



; NUMBER OF SEQ ID NOS: 133  
 ; SOFTWARE: CurationList version 0.1  
 ; SEQ ID NO 39  
 ; LENGTH: 1707  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: (5)..(1669)  
 US-10-262-445-39

Query Match 41.7%; Score 1423.8; DB 16; Length 1707;  
 Best local similarity 88.4%; Pred. No. 0;  
 Matches 1625; Conservative 0; Mismatches 82; Indels 132; Gaps 2;

QY	66	ACGGATGATGAACAAAGTTTACATCGGGAACCTTGAGCCCGCGCTGACCCCGCAGACCT	125
Db	1	ACGGATGATGAACAAAGTTTTCATCGGGAACCTGAGCCCGCGCTGACCCCGCAGACCT	60
QY	126	CCGGCAGCTCTTTGGGACAGGAGCTGCCCTGGCGGAGCAGGTCTGTGAAGTCGG	185
Db	61	CCGGCAGCTCTTTGGGACAGGAGCTGCCCTGGCGGAGCAGGTCTGTCAAGTCGG	120
QY	186	CTAGCGCTTGTGGACTACCCGACAGAACTGGGCCATCCGCGCATCGAGACCTCTC	245
Db	121	CTAGCGCTTGTGGACTACCCGACAGAACTGGGCCATCCGCGCATCGAGACCTCTC	180
QY	246	GGTAAAGTGAATTCATGCGGAATATGGAAGTGTGATCTCAAGTCTCTATAAAGCT	305
Db	181	GGTAAAGTGAATTCATGCGGAATATGGAAGTGTGATCTCAAGTCTCTATAAAGCT	240
QY	306	AAGGACGAGGAAATTCAGATTCGAAACATCCCTCTCTACCTGCGAGTGGAGTGTGA	365
Db	241	AAGGACGAGGAAATTCAGATTCGAAACATCCCTCTCTACCTGCGAGTGGAGTGTGA	300
QY	366	TGGAATTTGGCTCAATATGGACAGTGGAGATGTGGAACAGTCAACACAGACACAGA	425
Db	301	TGGAATTTGGCTCAATATGGACAGTGGAGATGTGGAACAGTCAACACAGACACAGA	360
QY	426	AACCGCGCTTTCACGTCATATGCAACAGAGAAAGCAAAATAGCCATGGAGAA	485
Db	361	AACCGCGCTTTCACGTCATATGCAACAGAGAAAGCAAAATAGCCATGGAGAA	420
QY	486	GCTAAGCGGGCATCAGTTTGAGAACTACTCTTCAAGATTTCTTACATCCCGGATGAA	545
Db	421	GCTAAGCGGGCATCAGTTTGAGAACTACTCTTCAAGATTTCTTACATCCCGGATGAA	480
QY	546	GCTGAGTCCCTTGGCCCTCTAGCGAGCCAGCGTGGGACCACTCTTCCCGGAGCA	605
Db	481	GCTGAGTCCCTTGGCCCTCTAGCGAGCCAGCGTGGGACCACTCTTCCCGGAGCA	540
QY	606	AGGCCAGCCCTGGGGGCACTTCTAGGCCAGACAGATTTCCCGCTGCGGATCCT	665
Db	541	AGGCCAGCCCTGGGGGCACTTCTAGGCCAGACAGATTTCCCGCTGCGGATCCT	600
QY	666	GGTCCCCACCCAGTTTCTGGTGCATCATCGGAAGAGGGCTTGACATTAAGAACAT	725
Db	601	GGTCCCCACCCAGTTTCTGGTGCATCATCGGAAGAGGGCTTGACATTAAGAACAT	660
QY	726	CATTAAGCAGACCAAGTCCCGGTAGATATCCATAGAAAAGAGAACTCTGGAGCTGAGA	785
Db	661	CATTAAGCAGACCAAGTCCCGGTAGATATCCATAGAAAAGAGAACTCTGGAGCTGAGA	720
QY	786	GAAGCTGTCCACATCCATCCCGACCCAGAGGGGACTTCTGAAGCATGCCGATGATCT	845
Db	721	GAAGCTGTCCACATCCATCCCGACCCAGAGGGGACTTCTGAAGCATGCCGATGATCT	780
QY	846	TGAAATCATGCAAAAGAGGCGAGATGAGACCAACTAGCCGAGAGATTCCTCTGAAAT	905
Db	781	TGAAATCATGCAAAAGAGGCGAGATGAGACCAACTAGCCGAGAGATTCCTCTGAAAT	840
QY	906	CTTGGCACAATGGCTTGGTGGAGACTGATTTGAAAAGAGGAGCAAAATTTGAAGAA	965

Db	841	CTTGGCCCACAATGGCTTGGTGGAGACTGATTTGAAAAGAGGACGAAATTTGAAGAA	900
QY	966	AATTGAACATGAAACAGGACCAAGATAAATCTCTCTTTTGGAGGATTTGAGCATATA	1025
Db	901	AAATGAACATGAAACAGGACCAAGATAAATCTCTCTTTTGGAGGATTTGAGCATATA	960
QY	1026	CAACCCGAAAGAACCATCTCTGTGAAGGGCACAGTCGAGGTCTGTGCCAGTGTGAGAT	1085
Db	961	CAACCCGAAAGAACCATCTCTGTGAAGGGCACAGTCGAGGTCTGTGCCAGTGTGAGAT	1020
QY	1086	AGAGATTATGAAGAAGCTGCTGAGGCTTTGAAAATGATATGCTGTGCTTTAACCAACA	1145
Db	1021	AGAGATTATGAAGAAGCTGCTGAGGCTTTGAAAATGATATGCTGTGCTTTAACCAACA	1074
QY	1146	AGCAATATGATCCCAAGGTTGAACCTCAGCGCACTTGGCATCTTTTCAACAGACTGTC	1205
Db	1075	-----	1074
QY	1206	CGTGTCTATCTCCACAGCAGGGCCCGGGAGCTCCCCCGTGTCCCTTACCACCCCTT	1265
Db	1075	-----	1074
QY	1266	CATACCACTCCGGATCTTCTCCAGCCTGTACCCCGATCACAGTTTGGCCCGTTCCT	1325
Db	1075	---TACCACTTCGGATCTTCTCCAGCCTGTACCCCGATCACAGTTTGGCCCGTTCCT	1131
QY	1326	GAATCATCTCTTATCCAGACAGGAGATTTGTAATCTTCTATCCCAACCCAGGCTGT	1385
Db	1132	GAATCATCTCTTATCCAGACAGGAGATTTGTAATCTTCTATCCCAACCCAGGCTGT	1191
QY	1386	GGCGCCCATCATCGGGAAGAGGGGACACATCAAAAGAGTGTGCGGAGATTCGCGGAGC	1445
Db	1192	GGCGCCCATCATCGGGAAGAGGGGACACATCAAAAGAGTGTGCGGAGATTCGCGGAGC	1251
QY	1446	CTTATCAAGATTGCCCTCGGAGAGCCAGCTCAGCGAAGAGTGTGTCATCATCAC	1505
Db	1252	CTTATCAAGATTGCCCTCGGAGAGCCAGCTCAGCGAAGAGTGTGTCATCATCAC	1308
QY	1506	CGGCGCACCGAAGCCAGTTTCAAGGCCCGAGGACCGATCTTTGGGAACTGAAAGAGA	1565
Db	1309	CTGCGCACCGAAGCCAGTTTCAAGGCCCGAGGACCGATCTTTGGGAACTGAAAGAGA	1368
QY	1566	AACTCTTTTAAACCCCAAGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA	1625
Db	1369	AACTCTTTTAAACCCCAAGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA	1428
QY	1626	CACAGCTGGCGGCTGATTGGCAAGGTGGCAAGACCGTGAACCACTGCGAAGCTTAA	1685
Db	1429	CACAGCTGGCGGCTGATTGGCAAGGTGGCAAGACCGTGAACCACTGCGAAGCTTAA	1488
QY	1686	CAGTGCAGAGTCTATCTGCTCTGTGACCAAAAGCCAGATGAAATGAGGAAGTATCGT	1745
Db	1489	CAGTGCAGAGTCTATCTGCTCTGTGACCAAAAGCCAGATGAAATGAGGAAGTATCGT	1548
QY	1746	CAGATTTATCGGCACCTTTCTTGTAGCCAGACTGCAACAGCGAAGATCAGGGAATTT	1805
Db	1549	CAGATTTATCGGCACCTTTCTTGTAGCCAGACTGCAACAGCGAAGATCAGGGAATTT	1608
QY	1806	ACAACAGGTGAAGCAGCAGGAGCAGAAATACCTCTCAGGAGTCCCTCAGCGCAGCAA	1865
Db	1609	ACAACAGGTGAAGCAGCAGGAGCAGAAATACCTCTCAGGAGTCCCTCAGCGCAGCAA	1668
QY	1866	GTGAGCTCCACAGCAGCAGCAACCAACCGATGAAAT	1904
Db	1669	GTGAGCTCCACAGCAGCAGCAACCAACCGATGAAAT	1707

RESULT 9  
 US-09-764-864-749  
 ; Sequence 749, Application US/09764864  
 ; Patent No. US20020132753A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Rosen et al.

Thu Jul 22 08:21:45 2004

```

; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT223
; CURRENT APPLICATION NUMBER: US/09/764,864
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 749
; LENGTH: 1186
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (511)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (774)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (777)
; OTHER INFORMATION: n equals a,t,g, or c
; US-09-764-864-749

Query Match      32.1%; Score 1096.8; DB 9; Length 1186;
Best Local Similarity 98.4%; Pred. No. 1.4e-290;
Matches 1145; Conservative 3; Mismatches 12; Indels 4; Gaps 4;

Qy 1376 CCCAGGCTGTGGCGCCATCATCGGGAAGAGGGGGACACATCAAAACAGCTGGCGAGAT 1435
Db 25 CCCACGCTGTGGCGCCATCATCGGGAAGAGGGGGACACATCAAAACAGCTGGCGAGAT 84
Qy 1436 TCGCGGAGCGCTCTATCAAGATTCGCCCTCGGGAGGCGCCAGAGCTCAGCGAAAGGATGG 1495
Db 85 TCGCGGAGCGCTCTATCAAGATTCGCCCTCGGGAGGCGCCAGAGCTCAGCGAAAGGATGG 144
Qy 1496 TCATCATCACCGGCGCCACCGGAAGCCAGTTCAGGCGCCAGGCGGATCTTTGGGAAC 1555
Db 145 TCATCATCACCGGCGCCACCGGAAGCCAGTTCAGGCGCCAGGCGGATCTTTGGGAAC 203
Qy 1556 TGAAGAGGAAACTTCTTTAAACCCAAAGAGAGTGAAGCTGGAAGCGCATATCAGAG 1615
Db 204 TGAAGAGGAAACTTCTTTAAACCCAAAGAGAGTGAAGCTGGAAGCGCATATCAGAG 263
Qy 1616 TGCCCTCTTCCACAGCTGCGCGGTGATTTGGCAAGGTGGCAAGACCGTGAACGAATGC 1675
Db 264 TGCCCTCTTCCACAGCTGCGCGGTGATTTGGCAAGGTGGCAAGACCGTGAACGAATGC 323
Qy 1676 AGAAGCTTAAACAGTGAGAGTCAATCGTCTGACCAACCGCCAGATGAAGATGAGG 1735
Db 324 AGAAGCTTAAACAGTGAGAGTCAATCGTCTGACCAACCGCCAGATGAAGATGAGG 383
Qy 1736 AAGTGAATCTCAGAAATATCGGCACTTTCTTGTAGCCAGATGACACGCGCAAGATCA 1795
Db 384 AAGTGAATCTCAGAAATATCGGCACTTTCTTGTAGCCAGATGACACGCGCAAGATCA 443
Qy 1796 GGGAAATTTG-TACAAAGTGTGAAGCAGGAGCAGAAATACCCCTCAGGAGTGGCTCA 1854
Db 444 GGGAAATTTG-TACAAAGTGTGAAGCAGGAGCAGAAATACCCCTCAGGAGTGGCTCA 503
Qy 1855 CAGCGCAGCAAGTGAAGTCTCCACAGCCACCAACCAACCGATGAATGTAGCCCTTC 1914
Db 504 CAGCGCA-NAAGTGAAGTCTCCACAGCCACCAACCAACCGATGAATGTAGCCCTTC 562
Qy 1915 CAACCTGACAGAAATGAGACCAAAACCGCAGCAGATCGGAGCAAAACCAAGACCA 1974
Db 563 CAACCTGACAGAAATGAGACCAAAACCGCAGCAGATCGGAGCAAAACCAAGACCA 622
Qy 1975 TCTGAGCAATGAGAAATCTGCGGAGGCGCCAGGAGTCTTCCGAGGCCCTTGAGAACCC 2034
Db 623 TCTGAGCAATGAGAAATCTGCGGAGGCGCCAGGAGTCTTCCGAGGCCCTTGAGAACCC 682
Qy 2035 AGGGCGCAGAGGGGGGAGAGGTGAGCCAGGTTTGCCAGAACCAACCGAGGATTCCTC 2094
Db 61 TGATTTCTTGAATCATGCGAAGAGCGAGATGAGACCAAACTAGCCGAGAGATTCCTC 120

683 AGGGCGCAGAGGGGGGAGAGGTGAGCCAGGTTTGCCAGAAACACCGAGGCCCGCT 742
2095 CCGCGCCCGCCAGGGCTTCTGAGGCTTTCAGGCAT-CCATTTCCACCATCCAGTGGATCTC 2153
743 CCGCGCCCGCCAGGGCTTCTGAGGCTTTCAGGCATTCAGCATCCATCCAGTGGATCTC 802
2154 TCTGAACTCCCGACGAGCTATCCCTTTTGTGAACTAACATAGGTGAACGTTTCAA 862
803 TCTGAACTCCCGACGAGCTATCCCTTTTGTGAACTAACATAGGTGAACGTTTCAA 862
2214 GCCAAGCAAAATGACACCCCTTTTCTGTGGCAAACTGCTCTGTACATGTGTACATA 2273
863 GCCAAGCAAAATGACACCCCTTTTCTGTGGCAAACTGCTCTGTACATGTGTACATA 922
2274 TTAGAAGGGAAGATGTTAAGATATCTGGCTCTGGTGTACAGAGGTGCTGAGCGGT 2333
923 TTAGAAGGGAAGATGTTAAGATATCTGGCTCTGGTGTACAGAGGTGCTGAGCGGT 982
2334 AATATATTTTAGAAATAATATATCAATAAATCAACTCACTCACTCACTCACTCACT 2393
983 AATATATTTTAGAAATAATATATCAATAAATCAACTCACTCACTCACTCACTCACT 1042
2394 TAATTTTTTTTCTTTTAAAGAGAGAGGAGTCTTTTCTAGACTTTTAAAGAAATAAGTCT 2453
1043 TAATTTTTTTTCTTTTAAAGAGAGAGGAGTCTTTTCTAGACTTTTAAAGAAATAAGTCT 1102
2454 TTGGAGGTCTCAGGTTGAGAGAGGAGTCTTTGAGGCGCCACCGCACAATAATTCACCCAGA 2513
1103 TTGGAGGTCTCAGGTTGAGAGAGGAGTCTTTGAGGCGCCACCGCACAATAATTCACCCAGA 1162
2514 GGGAAATCTGCTCGGAAGGACACT 2537
1163 GGGAAATCTGCTCGGAAGGACACT 1186

RESULT 10
US-09-764-864-330
; Sequence 330, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT223
; CURRENT APPLICATION NUMBER: US/09/764,864
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 330
; LENGTH: 822
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (441)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (455)
; OTHER INFORMATION: n equals a,t,g, or c
; US-09-764-864-330

Query Match      22.7%; Score 776.2; DB 9; Length 822;
Best Local Similarity 97.8%; Pred. No. 2.1e-202;
Matches 805; Conservative 10; Mismatches 5; Indels 3; Gaps 3;

Qy 779 CTCGAGAGAGCTGTCACTATCCATCCACCCAGAGGGAGCTTCTGAGCATGCGCA 838
Db 1 CTCGAGAGAGCTGTCACTATCCATCCATCCACCCAGAGGGAGCTTCTGAGCATGCGCA 60
Qy 839 TGATTTCTTGAATCATGCGAAGAGCGAGATGAGACCAAACTAGCCGAGAGATTCCTC 898
Db 61 TGATTTCTTGAATCATGCGAAGAGCGAGATGAGACCAAACTAGCCGAGAGATTCCTC 120
```





QY	1024	TACAACCCGGAAAGAACCATCACTGTGAAGGGCACAGTTGAGGCCTGTGCCAGTCTCTGAG	1083
Db	961	TATATCCGAGACGCACTATTACAGTTAAGGCAATGTTGACACATGTGCCAAAGCTGAG	1020
QY	1084	ATAGAGATTATGAAGAAGCTGGGTGAGGCTTTGAAATGATATGCTGGCTGTTAAACAA	1143
Db	1021	GAGGAGATCATGAAGAAATCAGGAGTCTTATGAAATGATTTGCTTCTATGAATCTT	1080
QY	1144	CAAGCCAACTGATCCAGGGTTGAACCTCAGCGCACTTGGCATCTTTTCAACAGAGCTG	1203
Db	1081	CAAGCACATTAATCTCTGATTAAATCTGAACGCTTGGGTCTGTTCCACCCACTTCA	1140
QY	1204	TCCGTGCTATCTCCACAGCAGGCGCCCGGGAGCTCCCCCGCTGCCCTTACCACCCC	1263
Db	1141	GGGATGCACCTCCCACTCAGGCGCCCTTCAGCCATGACTCCT	1185
QY	1264	TTCACTACCACATCCGGATATCTTCAGGCTGTACCCCATCACAGTTTGGCCGGTTC	1323
Db	1186	-----CCCTACCCGAGTTTGAGC-----	1204
QY	1324	CCGCATCATCACTCTTATCCAGACGAGGATTTGAACTCTTTCATCCCAACCCAGGCT	1383
Db	1205	-----AATCAGAAACGGAGACTGTTCACTCTGTTTATCCAGCTCTATCA	1248
QY	1384	GTGGGCGCCATCATCGGGAAGAAGGGGGCACATCAACAGCTGGCGAGATTGCGCGGA	1443
Db	1249	GTCCGTGCCATCATCGGCAAGCAGGGCCAGCACATCAAGCAGCTTCTCGCTTTGCTGGA	1308
QY	1444	GCCTCTATCAAGATTGCCCTGTCGGGAAGGCCAGACGTCAAGCAAGGATGGTCATCATC	1503
Db	1309	GCTTCAATTAAGATTGCTCCAGCGGAAGCACCCAGATGCTAAGTGAGGATGGGTATTATC	1368
QY	1504	ACCGGGCCACCGAAGCCAGTTCAAAGCCCCAGGACCGATCTTTTGGGAACTCAAGAG	1563
Db	1369	ACTTGACCACCAAGAGGCTCAGTTCAAGGCTCAGGGAAGATTATGGAATAATTAAAGAA	1428
QY	1564	GAATAACTTCTTTAAACCCCAAGAAGTGAAGCTTGAAGCGCATATCAGAGTGCCCTCT	1623
Db	1429	GAATACTTTGTAGTCTTAAGAGAGGTTGAACTTGAAGCTCATATCAGAGTGCCATCC	1488
QY	1624	TCCAAGTGGCCGGTGATTGGGCAAGGTGGCAAGCCGTGAACGACTGCAGAACTTA	1683
Db	1489	TTTGCTGTGCGAGAGTTATTGGAAAGAGGAGCAAAACGGTGAATGAATTCAGAATTG	1548
QY	1684	ACCAGTGCAGAGTCACTCGTGCCTCGTACCACAAACGCCAGATGAAATGAGGAAGTGATC	1743
Db	1549	TCAAGTGCAGAAGTTGTTGTCCCTCGTGAACAGACACCTGATGAGATGACCAAGTGTT	1608
QY	1744	GTCAGAATTTATCGGGCACTTTTTCGTAGCCAGACTGCACAGCGCAAGATCAGGGAATT	1803
Db	1609	GTCAAAATAACTGGTCACTTCTATGTTGCCAGTTGGCCACAGAAAAAATTCAGGAATT	1668
QY	1804	GTACAAACAGGTGAAGCAGAGCAGAGAATATACCTCAGGGAGTGCCTCAC	1855
Db	1669	CTGACTCAGGTAAGCAGACCAACACACAGAAGGCTCTGCAAGTGGACCAAC	1720

RESULT 12  
US-09-850-716A-347  
; Sequence 347, Application US/09850716A  
; Patent No. US20020115139A1  
; GENERAL INFORMATION:  
; APPLICANT: Kalos, Michael D.  
; APPLICANT: McNeill, Patricia D.  
; APPLICANT: Retter, Marc W.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER  
; FILE REFERENCE: 210121.455C15  
; CURRENT APPLICATION NUMBER: US/09/850,716A  
; CURRENT FILING DATE: 2001-05-07  
; NUMBER OF SEQ ID NOS: 440  
; SOFTWARE: FastSeq for Windows Version 3.0

Query ID	NO	347	1740	Score	712	DB	9	Length	1740								
1	ATGAA	CAAG	CTTTAC	AT	CGG	AA	CTGTAG	CG	CGCG	CGG	CTCAC	CGCG	CGAC	CGACT	CCG	CAG	132
1	ATGAA	CAAA	CTGTAT	AT	TCG	AA	AA	CT	CAG	CG	AA	CGCG	CGCC	CT	TCG	CA	60
133	CT	CTTT	GGG	CAC	AGG	AG	CT	CGCC	CT	GG	CGG	CAC	AGT	CT	CT	CT	192
61	AT	CT	TT	C	AAG	AC	CGC	AG	AT	CC	CGT	GT	CGG	AC	CT	CG	120
193	TT	CG	TG	GA	CT	AC	CC	CG	CA	AG	AA	CT	CGG	CA	AT	CG	252
121	TT	CG	TG	GA	CT	GC	CG	CA	AG	AG	CT	CGG	CG	CT	TT	CA	180
253	GT	GA	AT	TG	AT	CGG	AA	AA	TC	AT	CG	AG	TT	GA	TT	CA	312
181	AT	AG	AA	CT	GC	CGG	AA	CC	CA	TAG	AG	TT	GA	CA	CT	CG	240
313	AG	GA	AA	AT	TG	AT	CGG	AA	CT	CG	CG	CA	CT	CT	CT	CT	372
241	CG	GA	AA	CT	TG	AT	CGG	AA	CT	CG	CG	CA	CT	CT	CT	CT	300
373	TT	GG	CT	CA	AT	TG	GA	CT	CGG	AA	CT	CGG	CA	CT	CG	CA	432
301	CT	AG	TCC	AG	TAT	GG	AG	TG	GG	AG	CT	GT	GA	CA	CT	GA	360
433	GT	TC	CA	CG	TC	CA	TAT	GC	AA	GA	GA	GA	GA	GA	GA	GA	492
361	GT	TG	TA	AT	TG	TA	CT	AT	TC	AG	TA	AG	CA	CT	AG	CA	420
493	GG	CA	CT	CA	GT	TTG	AA	CT	ACT	CT	CT	CA	AG	TT	CC	TA	552
421	GG	AT	TT	CA	GT	TAG	AA	AT	TT	CA	CT	TC	AA	AG	TAG	CT	480
553	TC	CC	CT	TC	CG	CC	CT	C	AC	G	AG	CC	CG	GT	GG	CA	609
481	CAG	CA	AA	AC	CC	CT	TG	CA	CG	CC	CG	CG	CG	GT	GG	CA	540
610	-----	CA	CG	CC	CT	GG	GG	CA	CT	CT	C	AG	CC	CA	CT	CT	663
541	AG	GC	AG	GG	GT	CT	CC	AG	AT	CC	GT	AT	CC	AA	GC	AT	600
664	CT	GG	TC	CC	CA	CC	CG	AT	TT	GG	TC	CA	TC	CG	AA	AG	723
601	CT	GG	TT	CC	CA	CC	AA	TT	TT	GG	AG	CC	AT	C	A	G	660
724	AT	CA	CT	TA	GC	AC	CC	AG	TC	CG	GG	TAG	AT	T	CC	AT	783
661	AT	CA	CC	AA	CA	GC	CC	AG	TC	TA	AA	AT	CG	AT	GT	CC	720
784	GAG	AG	CC	CT	GT	TC	CA	AT	CC	AC	CC	CG	AG	GG	GA	CT	843
721	GAG	AG	TC	GA	TT	TA	CT	AT	CT	CT	CT	ACT	CT	CT	CT	CT	780
844	CT	TG	AA	AT	CA	TG	CAG	AA	G	AG	C	AG	AT	CG	CA	AT	903
781	CT	GA	GA	TT	AT	GC	AT	TA	AG	GA	CT	CA	GA	AT	TA	AA	840
904	AT	CT	TC	GC	CA	CA	AA	TT	GG	CT	TC	GT	GG	AG	AC	TG	963</





QY 724 ATCACTAAGCAGACCCAGTCCCGGTAGATATCCATAGAAAAGAGAACTCTGGAGCTGCA 783  
Db 661 ATCAACCAACAGACCCAGTCTAAATATGATGTCCACCGTAAAGAAATGCGGGGCTGCT 720  
QY 784 GAGAAGCCTGTCACCATCCATGCCACCCAGAGGGAGCTTCTGAAGCATGCCGATGATT 843  
Db 721 GAGAAGTCGATTACTATCTCTCTACTCTCTGAAGGCACCTCTCGGGCTGTAACTATT 780  
QY 844 CTTGAATCATGCAAGAGAGCAGATGAGACCAAACTAGCGAAGAGATTCCTCTGAAA 903  
Db 781 CTGGAGATTATGCAATAGGAGCTCAAGATATAAATTCACAGAGAGATCCCTCTGAAG 840  
QY 904 ATCTTGCACACAAATGCTTGGTTGGAAGCTGATTTGGAAGAAAGAGAGAAATTTCAAG 963  
Db 841 ATTTAGCTCATATAATCTTTGTTGGACGCTTATTGTTAAAGAGAGATCCCTCTGAAG 900  
QY 964 AAAATGAACATCAAAACAGGACCCAGATAACATCTCATCTTTGAGGATTTGAGATA 1023  
Db 901 AAAATTGAGCAGACACAGACACTAAATCAGATATCTCAATTGCAGGAATTGACGCTG 960  
QY 1024 TACAACCCGAAAGAACCACTACTGTGAAGGCACAGTTCAGGCCTGTGCGAGTGTGAG 1083  
Db 961 TATAATCCAGACGCACTATTACAGTTAAGGCAATGTTGAGACATGTGCCAAGCTGAG 1020  
QY 1084 ATAGAGATTATGAAGAGCTGCTGAGGCTTTTGAATAATGATATGCTGCTGTTAACCAA 1143  
Db 1021 GAGGAGATCATGAAGAAATCAGGAGTCTTATGAAATGATATGCTTCTATGATCTT 1080  
QY 1144 CAAGCAATCTGATCCAGGTTGAACCTCAGCGCACTTGGCATCTTTTCAACAGGACTG 1203  
Db 1081 CAAGCACATTTAATCTCTGATTAAATCTGAAGCCTTGGTCTGTTCACCACTTCA 1140  
QY 1204 TCCGTGCTATCTCCACAGCAGGCGCCCGGAGCTCCCGCGCTGCCCGCTACCAACCC 1263  
Db 1141 GGGATGCCACCTCCACCTCAGGCGCCCTTTCAGCATGACTCTT 1185  
QY 1264 TTCACCTACCACTCGGATACCTTCCAGCTGTACCCCATCACAGTTTGGCCGCTTC 1323  
Db 1186 -----CCCTACCGCAGTTTGAGC----- 1204  
QY 1324 CGGCATCACTACTTATCCAGAGGAGATTGTGAATCTTCTATCCCAACCCAGGCT 1383  
Db 1205 -----AATCAGAAACGGAGACTGTTTCATCTGTTTATCCAGCTCTATCA 1248  
QY 1384 GTGGGGCCATCATCGGGAAGAGGGGGCACATCAACACGCTGCGAGATTCCCGGA 1443  
Db 1249 GTGGGTGCCATCATCGGCAAGCAGGCGCCAGCACATCAAGCAGCTTCTCGCTTGTCTGA 1308  
QY 1444 GCCTCTATCAAGATTGCCCTCGGGAAGCCGAGCGTCAAGCAAGGATGTCATCATC 1503  
Db 1309 GCTTCAATTAAGATTGCTCCAGCGGAAGCACAGATGCTTAAAGTGAGGATGTTGATTATC 1368  
QY 1504 ACCGGGCCACCGGAAGCCAGTTCAAGGCCAGGACGATCTTTGGGAAACTGGAAGAG 1563  
Db 1369 ACTGGAACCAACAGAGGCTCAGTTCAAGGCTCAGGGAAGATTTATGGAAATTAAGAA 1428  
QY 1564 GAAACTCTTTTAAACCCCAAGAAAGTGAAGCTGGAAGGCGCATATCAGAGTGCCCTCT 1623  
Db 1429 GAAAACTTTGTTAGTCTTAAAGAGAGGTTGAACTTGAAGCTCATATCAGAGTGCCATCC 1488  
QY 1624 TCCACAGTGGCGGGTGATTGGCAAGGTGGCAAGCCGTGAACCTCAGAACTTCA 1683  
Db 1489 TTTGCTGTCGACAGATTATTTGAAAAAGGAGGCAAAAGCGTGAATGAACCTTCAAAATTTG 1548  
QY 1684 ACCAGTGCAGAGTCACTGTCCTCGTGACCAACCGCCAGATGAAATGAGGAGTGAATC 1743  
Db 1549 TCAAGTGCAGAGTTGTTGTCCTCTGTCAGCAGACACTGATGAGATGACCAAGTGGTT 1608  
QY 1744 GTCAAGATTATCGGGCACTTCTTTGCTAGCCAGACTGCACAGGCCAAGATCAGGGAAT 1803  
Db 1609 GTCAAAATAACTGTGTCATCTATGCTTGGCAGGTTGCCAGAGAAATTCAGAAAT 1668

QY 1804 GTACAACAGGTGAAGCAGGAGGAGCAGAAATATCCCTCAGGAGTGCCTCTAC 1855  
Db 1669 CTGACTCAGGTAAAGCAGCACCAACACAGAAAGCTCTGCAAGAGTGGACAC 1720  
RESULT 15  
US-10-117-982-347  
; Sequence 347, Application US/10117982  
; Publication No. US20030138438A1  
; GENERAL INFORMATION:  
; APPLICANT: FOY, Teresa M.  
; APPLICANT: Fauger, Gary R.  
; APPLICANT: Vedvick, Thomas S.  
; APPLICANT: Carter, Darrick  
; APPLICANT: Watanabe, Yoshihiro  
; APPLICANT: Henderson, Robert A.  
; APPLICANT: Kalos, Michael D.  
; APPLICANT: Mericle, Barbara  
; APPLICANT: Spies, Gregory A.  
; APPLICANT: Fan, Liqun  
; APPLICANT: Wang, Tongtong  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; FILE REFERENCE: 210121.455C18  
; CURRENT APPLICATION NUMBER: US/10/117,982  
; CURRENT FILING DATE: 2002-04-05  
; NUMBER OF SEQ ID NOS: 484  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 347  
; LENGTH: 1740  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-117-982-347

Query Match 20.9%; Score 712; DB 15; Length 1740;  
Best Local Similarity 64.5%; Pred. No. 1.8e-184;  
Matches 1156; Conservative 0; Mismatches 555; Indels 81; Gaps 3;

QY 73 ATGAACAGCTTTATCATCGGGAACCTGAGCCCGCGCTCACGCCGACGACCTCCGGCAG 132  
Db 1 ATGAACAAACTGTATATCGGAAACCTCAGCGAAGACGCGCGCCCTCGACCTAGAAAT 60  
QY 133 CTCCTTTGGGACAGCAAGCTGCCCTCGGGGAGCAGGTCTCTGCTGAAGTCCGGCTACGCC 192  
Db 61 ATCTTCAAGAGCGCCAGATCCCGGTGTCGGACCTTCTCTGTGAAGACTGCTACGCG 120  
QY 193 TTCGTGACTACCCCGACAGAACTGGGCCATCGCGCCATCGAGACCTCTCGGTAA 252  
Db 121 TTCGTGGACTGCCCGACGAGCTGGCCCTCAAGGCCATCGAGCGCTTTCAGGTAA 180  
QY 253 GTGGAATTGCATGGGAAATCATGGAAGTTGATTACTCAGTCTCTTAAAGCTAAGGAGC 312  
Db 181 ATGAACCTGACCGGAAACCCATAGAGTTGAGCTCGGTCCCAAAAGGCAAGGATT 240  
QY 313 AGGAAATTCAGATTTCGAAACATCCCTCTCACTGAGTGGAGGTGTTGATGAGCTT 372  
Db 241 CGGAACTTCAGATACGAAATATCCCGCTCATTTTACAGTGGGAGGTCTGGATAGTTA 300  
QY 373 TTGGCTCAATATGGGACAGTGGAGATGTGGAACAAGTCAACACAGACACAGAAACCGCC 432  
Db 301 CTAGTCAGTATGAGTGGTGGAGAGCTGTGAGCAAGTGAACATGACTCGAAACTGCA 360  
QY 433 GTTGTCACTCATATGCAACAGAGAGAGCAAAATAGCCATGGAGAGCTAAGC 492  
Db 361 GTTGAAATGTAACTTATCCAGTAAGGACCAAGCTAGACAACTAGACAACTGAT 420  
QY 493 GGGCATCAGTTTGAAGACTACTCTCTCAAGATTTCTTACATCCCGGATGAGAGGTGAGC 552  
Db 421 GGATTCAGTTAGAGATTTTCACTTGAAGTAGCTTATATCCCTGATGAACCGCGCC 480  
QY 553 TCCCTTCGCGCCCTCAGCGAGGCCAGCGTGGGACCACTCTTCCCGGAGCAAGGC--- 609  
Db 481 CAGCAAAACCCCTTGCAGCAGCCCGAGGTTCGCGGGGCTTGGGACAGAGGGCTCTCTCA 540

QY	610	-----CAGCGCCCTGGGGCACTTCTCAGGCCAGACAGATTGATTTCCGCTGGGATC	663
Db	541	AGCAGGGGTCTCCAGGATCCGATATCAAGAGAAACCATGTGATTTGCTCTGCGCCTG	600
QY	664	CTGGTCCCAACCCAGTTTGTGTGTCATATCGGAAGAGGGCTTGACCAATAAGAAC	723
Db	601	CTGGTCCCAACCCAAATTTGTGTGAGCATATAGGAAAGAGGTGCCACCATTCGGAAC	660
QY	724	ATCATTAGACAGACCCAGTCCCGGTAGATATCCATAGAAAGAGAACTCTGAGCTGCA	783
Db	661	ATCACCAGACAGACCCAGTCTAAATCGATGCCACCGTAAGAAATGCGGGGCTGCT	720
QY	784	GAGAAGCTCTCACCATCCATGCCACCCAGGGGACTTCTGAAGCATGCCCGATGAT	843
Db	721	GAGAGTCGATTACTTCTCTCTACTCTCTGAGGACCTCTGCGGCTTGAAGTCTATT	780
QY	844	CTTGAATCATGAGAAAGAGGAGATGAGACCAAACTAGCCGAGAGATTCCTCTGAAA	903
Db	781	CTGGAGATTATGATAGGAAGCTCAAGATATAAATTCACAGAAGAGATCCCTTGAAG	840
QY	904	ATCTTGGCACACATGCTTGTGTGGAAGTGAATGGAAGAAAGGAGAAATTTGAAG	963
Db	841	ATTTTAGCTCATTAATAACTTTGTGGAGCTTTATGTGTAAGAGGAAGAAATCTTAAA	900
QY	964	AAAAATTGAACATGAACAGGGACCAAGATAACAATCTCATCTTTGAGGATTTGAGCAT	1023
Db	901	AAATTCGACAGACACAGACACTAAAATCAGGATATCTCCATTCAGGAATTCAGCGTG	960
QY	1024	TACAAACCGGAAGAACCTCAGTGAAGGACAGTGGGCTGTCAGTGGCTGAG	1083
Db	961	TATAATCCAGAACCACTATTACAGTTAAAGGCAATGTTGAGATGTCGCGAAAGCTGAG	1020
QY	1084	ATAGAGATTATGAAGAGCTGCTGAGGCTTTGAAATGATATGCTGGCTGTGTTAACCA	1143
Db	1021	GAGGAGATCATGAGAAATCAGGAGTCTTATGAATGATATGCTTCTATGAATCTT	1080
QY	1144	CAAGCCATCTGATCCAGGTTGAACCTCAGCGCATTTGGCATCTTTTCAACAGGACTG	1203
Db	1081	CAAGCACATTTAAATTCCTGGATTAAATCTGAACGCTTTGGGTCTGTCTCCACCCACTTCA	1140
QY	1204	TCCGTGCTATCTCAGCAGGCGCCGCGGAGCTCCCGCTGCCCTACCAACCC	1263
Db	1141	GGGATGCCACCTCCACCTCAGGGCCCTTCAGCCATGACTCCCT	1185
QY	1264	TTCACTACCCACTCCGGATCTTCTCCAGCCTGTACCCCATCACAGTTTGGCCGCTTC	1323
Db	1186	-----CCCTACCGCAGTTTGAGC-----	1204
QY	1324	CGGCATCATCTTATCCAGAGAGGATGTGAATCTCTTATCCCAACCCAGGCT	1383
Db	1205	-----AATCAGAAACGGAGACTGTTCATCTGTTTATCCAGCTCTATCA	1248
QY	1384	GTGGGGCCCATCATCGGAGAGGGGGCACATCAACAGCTGGCGAGATTGCGCGGA	1443
Db	1249	GTGCGTGCCCATCATCGGAAGAGGGCCAGCATCAAGCAGCTTTCTCGTTGCTGGA	1308
QY	1444	GCCTCTATCAGATTGCCCTCGGAAGGCCAGAGCTCAGCGAAGAGATGGTCAATPC	1503
Db	1309	GCTTCATTAAGATTGCTCCAGGAGAGCACCAGATGCTTAAGTAGGATGGTGAATATC	1368
QY	1504	ACCGGCCCAACCGAAGCCAGTTCAAGGCCAGGACGGATCTTTGGGAACTGAAGAG	1563
Db	1369	ACTGGACCAACAGAGCTCAGTTCAAGGCTCAGGGAAGAAATTTATGGAATAATAAGAA	1428
QY	1564	GAATACTTTTAAACCCAAAGAGAGTGAAGCTGGAAGCATATCAGAGTGCCTCT	1623
Db	1429	GAATACTTTTAAAGAGAGAGTGAAGCTTTGAAGCTCATATCAGAGTGCCTATCC	1488
QY	1624	TCCACAGCTGGCGGGTATTTGCAAGAGTGGCAAGACCGTGAACCACTGCAAGAACTTA	1683
Db	1489	TTTGCTGCTGGCAGAGTTATTGGAAAGGAGGCAAAACGGTGAATGAATTCAGAAATTG	1548

Search completed: July 21, 2004, 13:09:48  
Job time : 1532 secs

QY	1684	ACCAAGTCAGAAAGTCACTGCTGCTGACCAAAACGCCAGATGAAATCAGGAAGTGATC	1743
Db	1549	TCAGTGCAGAAAGTTGTTGCTCCTGTCACCAAGACACCTGATGAGAAATGACCAAGTGGTT	1608
QY	1744	GTCAAGATTATCGGGCACTTCTTTGCTAGCCAGACTGCAACAGGCAAGATCAGGGAAT	1803
Db	1609	GTCAAAATAACTGTGTCACTTCTATGCTTCCAGGTTGCCCAAGAAAAATTCAGGAAAT	1668
QY	1804	GTACAACAGGTGAAGCAGCAGGAGCAGAAATACCTTCAGGGAGTCCGCTCAC	1855
Db	1669	CTGACTCAGGTAAGCAGCAGCAACCAACAGAGGCTCTGCAAGTGGACCAC	1720

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: July 16, 2004, 11:17:53 ; Search time 53.5 Seconds  
(without alignments)  
12269.364 Million cell updates/sec

Title: US-09-270-437D-6  
Perfect score: 1120  
Sequence: 1 ggcagcgaggagggcaggagga.....aaccttgaaatgtttattt 3412

Scoring table:

OLIGO  
Xgapop 60.0 , Xgapext 60.0  
Ygapop 60.0 , Ygapext 60.0  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 283366 seqs, 96191526 residues

Word size: 1

Total number of hits satisfying chosen parameters: 565818

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlp  
-Q=/cgn2.1/USPTO.spool p/US09270437/runat\_16072004.113128.13969/app\_query.fasta\_1.3591  
-DB=PIR 78 -QFWT=fastan -SURFIX=oligo.rpr -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=15 -MODE=LOCAL -OUTFWT=pt  
-NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09270437 @CGN 1.1.86 @runat\_16072004.113128.13969 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : PIR 78 : \*  
1: pir1 : \*  
2: pir2 : \*  
3: pir3 : \*  
4: pir4 : \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	9	0.8	319	2 F75420	hypothetical prote
2	9	0.8	366	2 AB1422	probable GTP-bind
3	9	0.8	366	2 AB1796	probable GTP-bind
4	9	0.8	398	1 WZBEN3	protein-serine/thr
5	9	0.8	399	2 JCT957	sex-determining re
C 6	9	0.8	431	2 T29850	hypothetical prote
C 7	9	0.8	491	1 FGHUB	fibrinogen beta ch
C 8	9	0.8	696	2 A28635	transcription fact
C 9	9	0.8	788	2 JS0747	regulatory protein
C 10	9	0.8	1226	2 T49915	pre-mRNA splicing
C 11	9	0.8	1234	2 S52099	phospholipase C be
12	9	0.8	2338	2 I73957	kinase-related pro
13	9	0.8	2347	1 TVHURS	kinase-related pro
C 14	8	0.7	63	2 T12121	NADH dehydrogenase

C 15	8	0.7	121	2 F72580	hypothetical prote
C 16	8	0.7	125	2 A71245	hypothetical prote
C 17	8	0.7	137	2 S37353	modulin (clone GmE
C 18	8	0.7	137	2 G75471	hypothetical prote
19	8	0.7	146	2 T16341	hypothetical prote
C 20	8	0.7	150	2 T17206	hypothetical prote
C 21	8	0.7	157	2 AE1293	shikimate kinase h
22	8	0.7	161	1 S76604	hypothetical prote
23	8	0.7	171	2 B89975	conserved hypothet
C 24	8	0.7	183	2 AF1883	hypothetical prote
25	8	0.7	202	2 T46586	ribosomal protein
C 26	8	0.7	205	2 T34724	probable membrane
C 27	8	0.7	208	2 T16953	hypothetical prote
C 28	8	0.7	219	2 A99194	iron (III) ABC tra
29	8	0.7	220	2 A36298	proline-rich prote
30	8	0.7	242	2 G96994	glycerol uptake fa
C 31	8	0.7	257	2 C96994	transcription regu
32	8	0.7	267	2 D83113	probable permease
33	8	0.7	270	2 AF1815	superoxide dismuta
34	8	0.7	279	2 D82281	ferric vibriobacti
35	8	0.7	285	2 H70781	hypothetical prote
36	8	0.7	301	2 G83182	hypothetical prote
37	8	0.7	305	2 T20906	interferon respons
38	8	0.7	311	2 A38558	hypothetical prote
39	8	0.7	315	2 A84634	transaldolase - He
C 40	8	0.7	316	2 E71812	hypothetical prote
C 41	8	0.7	320	2 AF1892	[NiFe] uptake hydr
42	8	0.7	324	2 G90896	probable transcrip
43	8	0.7	326	2 S56534	hypothetical 36.9K
44	8	0.7	326	2 D91287	hypothetical prote
45	8	0.7	326	2 G86128	hypothetical prote

## ALIGNMENTS

### RESULT 1

F75420  
Hypothetical protein - Deinococcus radiodurans (strain R1)  
C:Species: Deinococcus radiodurans  
C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 31-Mar-2000  
C:Accession: F75420  
R:White, O.; Eissen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;  
M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; M  
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.  
Science 286, 1571-1577, 1999  
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.  
A:Reference number: A75250; MUID:20036896; PMID:10567266  
A:Accession: F75420  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-319 <WHI>  
A:Cross-references: GB:AE001971; GB:AE000513; NID:G6458972; PIDN:AAF10810.1; PID:G64589  
A:Experimental source: strain R1  
C:Genetics:  
A:Gene: DR1229  
A:Map position: 1

Alignment Scores:			
Pred. No.:	11.7	Length:	319
Score:	9.00	Matches:	9
Percent Similarity:	100.00%	Conservative:	0
Best local Similarity:	100.00%	Mismatches:	0
Query Match:	0.82%	Indels:	0
DB:	2	Gaps:	0

US-09-270-437D-6 (1-3412) x F75420 (1-319)  
QY 1079 GCATGGCAGCGGCTCAACTGTGCC 1053  
Db 7 AlaLeuAlaGlnAlaSerThrValPro 15  
RESULT 2  
AB1422



Thu Jul 22 08:21:43 2004

us-09-270-437d-6.Oligo.rpr

```

probable GTP-binding protein homolog lmo2779 [imported] - Listeria monocytogenes (strain
C;Species: Listeria monocytogenes
C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001
C;Accession: AB1422
R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker
D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma
ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,
A;Title: Comparative genomics of Listeria species.
A;Reference number: AB1077; MUID:21537279; PMID:11679669
A;Accession: AB1422
A;Status: preliminary
A;Molecule type: DNA
A;Cross-references: GB:NC_003210; PIDN:CAD00992.1; PID:g16412279; GSPDB:GN00177
A;Experimental source: strain EGD-e
C;Genetics:
C;Superfamily: yeast probable purine nucleotide-binding protein YBR025c
C;Superfamily: yeast probable purine nucleotide-binding protein YBR025c

Alignment Scores:
Pred. No.: 11.5 Length: 366
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.80% Indels: 0
DB: 2 Gaps: 0

US-09-270-437D-6 (1-3412) x AB1422 (1-366)
QY 1099 AAGCTGCGTGAGCGCCCTTGAATGAT 1125
DB 167 LysLeuArgGluAlaPheGluAsnAsp 175

RESULT 3
AH1796
Probable GTP-binding protein homolog lin2919 [imported] - Listeria innocua (strain Cliph
C;Species: Listeria innocua
C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001
C;Accession: AH1796
R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker
D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma
ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,
A;Title: Comparative genomics of Listeria species.
A;Reference number: AB1077; MUID:21537279; PMID:11679669
A;Accession: AH1796
A;Status: preliminary
A;Molecule type: DNA
A;Cross-references: GB:AL592022; PIDN:CA98144.1; PID:g16415460; GSPDB:GN00178
A;Experimental source: strain Cliph1262
C;Genetics:
C;Superfamily: yeast probable purine nucleotide-binding protein YBR025c
C;Superfamily: yeast probable purine nucleotide-binding protein YBR025c

Alignment Scores:
Pred. No.: 11.5 Length: 366
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.80% Indels: 0
DB: 2 Gaps: 0

US-09-270-437D-6 (1-3412) x AH1796 (1-366)
QY 1099 AAGCTGCGTGAGCGCCCTTGAATGAT 1125
DB 167 LysLeuArgGluAlaPheGluAsnAsp 175

```

```

RESULT 4
WZBEN3
protein-serine/threonine kinase (EC 2.7.1.1) - suid herpesvirus 1 (strain NIA-3)
N;Alternate names: ULL3 protein
C;Species: suid herpesvirus 1
C;Date: 30-Jun-1993 #sequence_revision 30-Jun-1993 #text_change 11-Jun-1999
C;Accession: B42744
R;de Wind, N.; Domen, J.; Berns, A.
J. Virol. 66, 5200-5209, 1992
A;Title: Herpesviruses encode an unusual protein-serine/threonine kinase which is nones
A;Reference number: A42744; MUID:92365105; PMID:1323689
A;Accession: B42744
A;Molecule type: DNA
A;Residues: 1-398 <DEW>
A;Cross-references: GB:M94870; NID:g334092; PIDN:AAA47481.1; PID:g334094
C;Genetics:
C;Superfamily: herpesvirus protein-serine/threonine kinase; protein kinase homology
C;Keywords: ATP; phosphotransferase; serine/threonine-specific protein kinase
F;78-378/Domain: protein kinase homology <KIN>
F;86-93/Region: protein kinase ATP-binding motif
F;103/Active site: Lys #status predicted

Alignment Scores:
Pred. No.: 11.4 Length: 398
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.80% Indels: 0
DB: 1 Gaps: 0

US-09-270-437D-6 (1-3412) x WZBEN3 (1-398)
QY 2 GCAGCGGAGGCGGAGGAGCGCGGG 28
DB 119 AlaAlaGluGluAlaArgSerAlaGly 127

RESULT 5
JC7957
sex-determining region Y (SRY)-like HMG-box protein 17, SOX17 protein - rice field eel
C;Species: Monopterus albus (rice field eel)
C;Date: 22-Jun-2003 #sequence_revision 22-Jun-2003 #text_change 07-Jul-2003
C;Accession: JC7957
R;Wang, R.; Cheng, H.; Xia, L.; Guo, Y.; Huang, X.; Zhou, R.
Biochem. Biophys. Res. Commun. 303, 452-457, 2003
A;Title: Molecular cloning and expression of Sox17 in gonads during sex reversal in the
A;Reference number: JC7957; MUID:22546426; PMID:12659838
A;Accession: JC7957
A;Molecule type: mRNA
A;Residues: 1-399 <WAN>
A;Cross-references: GB:AY100695
C;Comment: This protein functions as a transcriptional activator during spermatogenesis
ment.
C;Genetics:
A;Gene: Sox17
A;Map position: 5
C;Keywords: gonad differentiation; HMG box; sexual differentiation; Sox17; spermatogen

Alignment Scores:
Pred. No.: 11.4 Length: 399
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.80% Indels: 0
DB: 2 Gaps: 0

US-09-270-437D-6 (1-3412) x JC7957 (1-399)
QY 1265 TCACCTACCCACTCCGGATACCTCCCA 1291
DB 210 SerLeuProThrProAspThrSerPro 218

```

RESULT 6

T29850

hypothetical protein C49C8.5 - *Caenorhabditis elegans*

C/Species: *Caenorhabditis elegans*

C/Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999

C/Accession: T29850

R/Johnson, D.; Bradshaw, H.

submitted to the EMBL Data Library, June 1996

A/Description: The sequence of *C. elegans* cosmid C49C8.

A/Reference number: Z20698

A/Accession: T29850

A/Status: preliminary; translated from GB/EMBL/DBD

A/Molecule type: DNA

A/Residues: 1-431 <JOH>

A/Cross-references: EMBL:U61945; PIDN:AA03126.1; GSPDB:GN00022; CESP:C49C8.5

A/Experimental source: strain Bristol N2; clone C49C8

C/Genetics:

A/Gene: CESP:C49C8.5

A/Map position: 4

A/Introns: 17/3; 66/2; 106/3; 151/2; 187/1; 233/2; 302/3; 329/1; 351/2; 377/2

Alignment Scores:			
Pred. No.:	11-3	Length:	431
Score:	9.00	Matches:	9
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	0.82%	Indels:	0
DB:	2	Gaps:	0

US-09-270-437D-6 (1-3412) x T29850 (1-431)

Qy 2590 AAGGAGACGGTATCCCTTCGTGACA 2564

|||||

Db 3 LysGluThrValSerLeuLeuLeuThr 11

RESULT 7

FGHUB

fibrinogen beta chain precursor [validated] - human

N/Alternate names: coagulation factor I

N/Contains: fibrinopeptide B

C/Species: Homo sapiens (man)

C/Date: 24-Apr-1984 #sequence\_revision 31-Mar-1993 #text\_change 08-Dec-2000

C/Accession: B43568; A90469; B90469; I37389; A94433; A90437; A94309; G54223; A03121; B37

R/Chung, D.W.; Harris, J.E.; Davie, E.W.

Adv. Exp. Med. Biol. 281, 39-48, 1990

A/Title: Nucleotide sequences of the three genes coding for human fibrinogen.

A/Reference number: A43568; MUID:91344740; PMID:2102623

A/Accession: B43568

A/Molecule type: DNA

A/Residues: 9-191, 'P', 193-491 <CHU>

R/Chung, D.W.; Que, B.G.; Rixon, M.W.; Mace Jr., M.; Davie, E.W.

Biochemistry 22, 3244-3250, 1983

A/Title: Characterization of complementary deoxyribonucleic acid and genomic deoxyribon

A/Reference number: A90469; MUID:83283433; PMID:6688356

A/Accession: A90469

A/Molecule type: DNA

A/Residues: 1-38 <CH1>

A/Accession: B90469

A/Molecule type: mRNA

A/Residues: 9-191, 'A', 193-491 <CH2>

A/Cross-references: GB:J00129; NID:g182429; PIDN:AA52429.1; PID:g182430

R/Huber, P.; Dalton, J.; Courtois, G.; Laurent, M.; Assouline, Z.; Marguerie, G.

Nucleic Acids Res. 15, 1615-1625, 1987

A/Title: Characterization of the 5'-flanking region for the human fibrinogen beta gene.

A/Reference number: I37389; MUID:87146483; PMID:3029722

A/Accession: I37389

A/Status: translated from GB/EMBL/DBD

A/Molecule type: DNA

A/Residues: 1-38 <HUB>

A/Cross-references: EMBL:X05018; NID:g31400; PIDN:CAA28674.1; PID:g31401

R/Henschen, A.; Lottspeich, F.; Southan, C.; Topfer-Petersen, E.

In Protides of the Biological Fluids, Proc. 28th Colloq., Peeters, H., ed., pp.51-56, Pe

A/Title: Human fibrinogen: sequence, sulfur bridges, glycosylation and some structural v

A:Reference number: A94433  
A:Contents: carbohydrate binding  
A:Accession: A94433  
A:Molecule type: protein  
A:Residues: 31-137,'QS',140-144,'QF',147-491 <HEN>  
R:Watt, K.W.K.; Takagi, T.; Doolittle, R.F.  
Biochemistry 18, 68-76, 1979  
A:Title: Amino acid sequence of the beta chain of human fibrinogen.  
A:Reference number: A90437; MUID:79124640; PMID:420779  
A:Accession: A90437  
A:Molecule type: protein  
A:Residues: 31-144,'QF',147-231,'D',233-330,'E',332-491 <WAT>  
R:Blomback, B.; Hessel, B.; Hogg, D.  
Thromb. Res. 8, 639-658, 1976  
A:Title: Disulfide bridges in NH-2-terminal part of human fibrinogen.  
A:Reference number: A94309; MUID:76225080; PMID:936108  
A:Contents: disulfide bonds  
A:Accession: A94309  
A:Molecule type: protein  
A:Residues: 31-112,'E',114-137,'QS',140-144,'QF',147-148 <BLO>  
R:Kunitake, S.T.; Carilli, C.T.; Lau, K.; Protter, A.A.; Naya-Vigne, J.; Kane, J.P.  
Biochemistry 33, 1988-1993, 1994  
A:Title: Identification of proteins associated with apolipoprotein A-I-containing lipoproteins  
A:Reference number: A54223; MUID:94162201; PMID:8117655  
A:Accession: G54223  
A:Molecule type: protein  
A:Residues: 164-174 <KUN>  
A:Note: identification of tryptic peptides from high-density lipoproteins  
R:Henschen, A.; Lottspeich, F.; Kehl, M.; Southan, C.  
Ann. N. Y. Acad. Sci. 408, 28-43, 1983  
A:Title: Covalent structure of fibrinogen.  
A:Reference number: A90037; MUID:83254370; PMID:6575689  
A:Contents: annotation; review, disulfide bonds  
R:Gardlund, B.; Hessel, B.; Marguerie, G.; Murano, G.; Blomback, B.  
Eur. J. Biochem. 77, 595-610, 1977  
A:Title: Primary structure of human fibrinogen. Characterization of disulfide-containing  
A:Reference number: A91249; MUID:77245999; PMID:891553  
A:Contents: annotation; disulfide bonds  
R:Doolittle, R.F.; Takagi, T.; Watt, K.; Bouma III, H.; Cottrell, B.A.; Cassman, K.G.; Folini, M.; Ottensen, S.; Magnusson, S.; Magnusson, S.; Ottensen, M., Folini, M.  
in Regulatory Proteolytic Enzymes and Their Inhibitors, Magnusson, S., Ottensen, M., Folini, M., Eds.  
A:Title: The structures of fibrinogen and fibrin.  
A:Reference number: A94437  
A:Contents: annotation; disulfide bonds  
R:Doolittle, R.F.  
Annu. Rev. Biochem. 53, 195-229, 1984  
A:Title: Fibrinogen and fibrin.  
A:Reference number: A90041; MUID:84305751; PMID:6383194  
A:Contents: annotation; review, EM structure, polymerization, ligands  
R:Chung, D.W.; Rixon, M.W.; Que, B.G.; Davie, E.W.  
Ann. N. Y. Acad. Sci. 408, 449-456, 1983  
A:Title: Cloning of fibrinogen genes and their cDNA.  
A:Reference number: A90038; MUID:83254384; PMID:6575700  
A:Contents: annotation  
R:Kirschbaum, N.E.; Budzynski, A.Z.  
J. Biol. Chem. 265, 13669-13676, 1990  
A:Title: A unique proteolytic fragment of human fibrinogen containing the Aalpha COOH-terminus  
A:Reference number: A37117; MUID:90337977; PMID:2143188  
A:Contents: annotation; hementin cleavage site  
A:Note: hementin, a protease from Haemeteria ghilianii, the giant South American leech, cleaves the COOH-terminus of fibrinogen to form a fragment containing the Aalpha COOH-terminus. This fragment is responsible for the formation of the soft clot.  
C:Comment: The soft clot is converted into the hard clot by factor XIIIa (fibrin-stabilizing factor) and between alpha chains (weaker) of different monomers.  
C:Comment: All fibrinogen chains are synthesized in the liver.  
C:Genetics:  
A:Gene: GDB:FGB  
A:Cross-references: GDB:119130; OMIM:134830  
A:Map position: 4q28-4q28  
A:Introns: 38/3; 102/3; 164/1; 240/1; 278/1; 320/1; 415/2  
C:Complex: The fibrinogen molecule is a hexamer containing two sets of alpha (see PIR:PC000000) and two sets of beta (see PIR:PC000000) chains. The alpha chains are contained in the core. Two three-chain coiled coils emerge from this core and contain the distal domain nodes.  
C:Function:

A;Description: fibrinogen cleaved by thrombin yields monomers that are polymerized into  
A;Pathway: blood coagulation  
C;Superfamily: fibrinogen beta chain; fibrinogen beta/gamma homology; fibrinogen disulfide  
C;Keywords: blood coagulation; coiled coil; glycoprotein; liver; plasma; pyroglutamic acid  
F;1-30/Domain: (or 4-30 or 15-30) signal sequence #status predicted <SIG>  
F;31-491/Product: fibrinogen beta chain #status experimental <MAT>  
F;31-44/Product: fibrinopeptide B #status experimental <APT>  
F;45-491/Product: fibrin beta chain #status experimental <FCB>  
F;45-47/Region: polymerization site  
F;99-228/Domain: fibrinogen disulfide ring homology <FDR>  
F;238-487/Domain: fibrinogen beta/gamma homology <FBG>  
F;31/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status experimental  
F;44-45/Cleavage site: Arg-Gly (thrombin) #status experimental  
F;95/Disulfide bonds: interchain (to alpha-55) #status experimental  
F;106/Disulfide bonds: interchain (to alpha-68) #status experimental  
F;110/Disulfide bonds: interchain (to gamma-45) #status experimental  
F;223/Disulfide bonds: interchain (to alpha-184) #status experimental  
F;227/Disulfide bonds: interchain (to gamma-161) #status experimental  
F;231-316,241-270,424-437/Disulfide bonds: #status experimental  
F;394/Binding site: carboxylate (Asn) (covalent) #status experimental

Alignment Scores:  
Pred. No.: 11.1 Length: 491  
Score: 9.00 Matches: 9  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 0.82% Indels: 0  
DB: 1 Gaps: 0

US-09-270-437D-6 (1-3412) x FGHUB (1-491)

QY 3171 CTTCTCTCGTGTTTTCGTAA 3145

Db 21 LeuLeuLeuCysValPheLeuValLys 29

RESULT 8

A29635  
transcription factor Sp1 - human (fragment)  
N;Alternate names: finger protein ZNF76  
C;Species: Homo sapiens (man)  
C;Date: 31-Mar-1989 #sequence\_revision 31-Mar-1989 #text\_change 05-Nov-1999  
C;Accession: A29635; G44256  
R;Kadonaga, J.T.; Garner, K.R.; Masiarz, F.R.; Tjian, R.  
Cell 51, 1079-1090, 1987  
A;Title: Isolation of cDNA encoding transcription factor Sp1 and functional analysis of  
A;Reference number: A29635; MUID:88080466; PMID:3319186  
A;Accession: A29635  
A;Molecule type: mRNA  
A;Residues: 1-696 <KAD>  
R;Cross-references: GB:J03133; NID:G339517; PIDN:AAA61154.1; PID:G339518  
R;Ragoussis, J.; Senger, G.; Mockridge, I.; Sansseau, P.; Ruddy, S.; Dudley, K.; Sheer, D.  
Genomics 14, 673-679, 1992  
A;Title: A testis-expressed Zn finger gene (ZNF76) in human 6p21.3 centromeric to the MH  
A;Reference number: A44256; MUID:93052398; PMID:1427894  
A;Accession: G44256  
A;Status: preliminary; not compared with conceptual translation  
A;Molecule type: nucleic acid  
A;Residues: 569-598 <RAG>  
A;Experimental source: T-cell line CEM  
A;Note: sequence extracted from NCBI backbone (NCBIP:125980)  
C;Genetics:  
A;Gene: GDB:SP1  
A;Cross-references: GDB:127453; OMIM:189906  
A;Map position: 19q13.1-19q13.3  
C;Keywords: DNA binding; transcription regulation; zinc finger

Alignment Scores:  
Pred. No.: 10.6 Length: 696  
Score: 9.00 Matches: 9  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 0.82% Indels: 0  
DB: 2 Gaps: 0

US-09-270-437D-6 (1-3412) x A29635 (1-696)

QY 893 ATCTCTCGGTAGTTTGTCATCT 867

Db 216 IleSerSerAlaSerLeuValSerSer 224

RESULT 9

JS0747  
regulatory protein Sp1 - rat  
C;Species: Rattus norvegicus (Norway rat)  
C;Date: 04-Dec-1992 #sequence\_revision 04-Dec-1992 #text\_change 01-Dec-2000  
C;Accession: JS0747; S25287  
R;Imataka, H.; Sogawa, K.; Yasumoto, K.; Kikuchi, Y.; Sasano, K.; Kobayashi, A.; Hayami  
submitted to JIPID, September 1992  
A;Reference number: JS0747  
A;Accession: JS0747  
A;Molecule type: mRNA  
A;Residues: 1-788 <IMA>  
R;Imataka, H.; Sogawa, K.; Yasumoto, K.; Kikuchi, Y.; Sasano, K.; Kobayashi, A.; Hayami  
EMBO J. 11, 3663-3671, 1992  
A;Title: Two regulatory proteins that bind to the basic transcription element (BTE), a  
A;Reference number: S25287; MUID:93010958; PMID:1356762  
A;Accession: S25287  
A;Molecule type: mRNA  
A;Residues: 1-122,'L',124-311,'A',313-788 <IM2>  
C;Keywords: DNA binding; transcription regulation

Alignment Scores:  
Pred. No.: 10.4 Length: 788  
Score: 9.00 Matches: 9  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 0.82% Indels: 0  
DB: 2 Gaps: 0

US-09-270-437D-6 (1-3412) x JS0747 (1-788)

QY 893 ATCTCTCGGTAGTTTGTCATCT 867

Db 308 IleSerSerAlaSerLeuValSerSer 316

RESULT 10

T49915  
pre-mRNA splicing factor ATP-dependent RNA helicase-like protein - Arabidopsis thaliana  
N;Alternate names: protein T24H18.180  
C;Species: Arabidopsis thaliana (mouse-ear cress)  
C;Date: 02-Jun-2000 #sequence\_revision 02-Jun-2000 #text\_change 02-Jun-2000  
C;Accession: T49915  
R;Bevan, M.; Robben, J.; Grymonprez, B.; Volckaert, G.; Bancroft, I.; Mewes, H.W.; Rueda  
submitted to the Protein Sequence Database, April 2000  
A;Reference number: Z25024  
A;Accession: T49915  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-1226 <BEV>  
A;Cross-references: EMBL:AL353013; GSPDB:GN00063; ATSP:T24H18.180  
A;Experimental source: cultivar Columbia; BAC clone T24H18  
C;Genetics:  
A;Map position: 5  
A;Introns: 2/3; 40/1; 111/1; 123/3; 183/3; 231/1; 279/3; 313/2; 349/3; 409/1; 485/3; 505/1  
Alignment Scores:  
Pred. No.: 9.8 Length: 1226  
Score: 9.00 Matches: 9  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 0.82% Indels: 0  
DB: 2 Gaps: 0

US-09-270-437D-6 (1-3412) x T49915 (1-1226)

Qy 499 GATGCCCGCTTAGCTTCTCCATGGCTA 473  
|||||  
Db 216 AspAlaArgLeuAlaSerProTrpLeu 224

RESULT 11  
S52099  
Phospholipase C beta 3 - human  
N;Alternate names: phospholipase c beta-3, phosphoinositide-specific  
C;Species: Homo sapiens (man)  
C;Date: 15-Jul-1995 #sequence revision 21-Jul-1995 #text\_change 17-Mar-1999  
A;Accession: S52099; A56833; A56854  
R;Lagercrantz, J.; Carson, E.; Phelan, C.; Grimmond, S.; Rosen, A.; Dare, E.; Nordenskjöld  
submitted to the EMBL Data Library, September 1994  
A;Description: Genomic organization and complete cDNA sequence of the human phosphoinosi  
A;Reference number: S52099  
A;Accession: S52099  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-1234 <LAG>  
A;Cross-references: EMBL:Z37573  
R;Lagercrantz, J.; Carson, E.; Phelan, C.; Grimmond, S.; Rosen, A.; Dare, E.; Nordenskjöld  
Genomics 26, 467-472, 1995  
A;Title: Genomic organization and complete cDNA sequence of the human phosphoinositide-s  
A;Reference number: A56833; MUID:95331781; PMID:7607669  
A;Accession: A56833  
A;Molecule type: DNA  
A;Residues: 1-200 <LA2>  
A;Cross-references: GB:Z37544  
R;Sinke, R.J.; Geurts van Kessel, A.  
Genomics 25, 568-569, 1995  
A;Title: Localization of the human phosphatidylinositol-specific phospholipase C beta-3  
A;Reference number: A56854; MUID:95309927; PMID:7789993  
A;Accession: A56854  
A;Status: preliminary; not compared with conceptual translation  
A;Molecule type: mRNA  
A;Residues: 1-193 <SIN>  
C;Genetics:  
A;Gene: GDB:PLCB3  
A;Cross-references: GDB:386061; OMIM:600230  
A;Map position: 11q13-11q13  
A;Introns: 33/3; 59/3; 82/3; 129/3; 156/2; 174/2; 199/3; 233/2; 288/3; 338/1; 418/2; 446  
1119/2; 1138 /3; 1167/3  
C;Superfamily: 1-phosphatidylinositol-4,5-bisphosphate phosphodiesterase I; 1-phosphatid  
odiesterase domain Y homology  
F;319-468/Domain: 1-phosphatidylinositol-4,5-bisphosphate phosphodiesterase domain X hom  
F;589-709/Domain: 1-phosphatidylinositol-4,5-bisphosphate phosphodiesterase domain Y hom

Alignment Scores:  
Pred. No.: 9.79 Length: 1234  
Score: 9.00 Matches: 9  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 0.82% Indels: 0  
DB: 2 Gaps: 0

US-09-270-437D-6 (1-3412) x S52099 (1-1234)

Qy 217 AGTTCTGTCGGGTAGTCACGAAGG 191  
|||||  
Db 1099 SerSerGlyArgGlySerProArgArg 1107

RESULT 12  
I73957  
kinase-related protein c-ros-1 precursor - rat  
N;Contains: protein-tyrosine kinase (EC 2.7.1.112) ros-1  
C;Species: Rattus norvegicus (Norway rat)  
C;Date: 02-Aug-1996 #sequence revision 02-Aug-1996 #text\_change 02-Jun-2000  
C;Accession: I73957; I56752; I73956  
R;Matsushima, H.; Shibuya, M.  
J. Virol. 64, 2117-2125, 1990  
A;Title: Tissue-specific expression of rat c-ros-1 gene and partial structural similarit  
A;Reference number: I56752; MUID:90219211; PMID:2139140

A;Accession: I73957  
A;Status: preliminary; translated from GB/EMBL/DBDJ  
A;Molecule type: mRNA  
A;Residues: 1-2338 <RES>  
A;Cross-references: GB:M35106; NID:g203599; PIDN:AAA40968.1; PID:g203600  
A;Accession: I56752  
A;Status: preliminary; translated from GB/EMBL/DBDJ  
A;Molecule type: mRNA  
A;Residues: 1-430,452-2338 <RE2>  
A;Cross-references: GB:M35104; NID:g203595; PIDN:AAA40966.1; PID:g203596  
A;Accession: I73956  
A;Status: preliminary; translated from GB/EMBL/DBDJ  
A;Molecule type: mRNA  
A;Residues: 1-430,452-1872, 'AC', 1875 <RE3>  
A;Cross-references: GB:M35105; NID:g203597; PIDN:AAA40967.1; PID:g203598  
C;Superfamily: kinase-related protein ros; LDL receptor YWTD-containing repeat homology  
C;Keywords: alternative splicing; ATP; autophosphorylation; glycoprotein; kinase-relate  
ific protein kinase  
F;753-793/Domain: LDL receptor YWTD-containing repeat homology <YW3>  
F;1335-2214/Domain: protein kinase homology <KIN>  
F;1943-1951/Region: protein kinase ATP-binding motif

Alignment Scores:  
Pred. No.: 8.98 Length: 2338  
Score: 9.00 Matches: 9  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 0.80% Indels: 0  
DB: 2 Gaps: 0

US-09-270-437D-6 (1-3412) x I73957 (1-2338)

Qy 569 AGCGAGCCGCGTGGGACACACTCTT 595  
|||||  
Db 656 SerGluProSerValGlyThrThrLeu 664

RESULT 13  
TYHURS  
kinase-related protein ros-1 precursor - human  
N;Alternate names: protein-tyrosine kinase mcf3 (activated ros-1)  
N;Contains: protein-tyrosine kinase (EC 2.7.1.112) ros-1  
C;Species: Homo sapiens (man)  
C;Date: 31-Dec-1988 #sequence revision 07-Oct-1994 #text\_change 11-Jun-1999  
A;Accession: A35512; A24223; A24421; A33081  
R;Birchmeier, C.; O'Neill, K.; Riggs, M.; Wigler, M.  
Proc. Natl. Acad. Sci. U.S.A. 87, 4799-4803, 1990  
A;Title: Characterization of ROS1 cDNA from a human glioblastoma cell line.  
A;Reference number: A35512; MUID:90280463; PMID:2352949  
A;Accession: A35512  
A;Molecule type: mRNA  
A;Residues: 1-2212, 'N', 2214-2227, 'QC', 2229-2347 <BIR>  
A;Cross-references: GB:M34353  
A;Experimental source: glioblastoma cell line SW-1088  
R;Matsushima, H.; Wang, L.H.; Shibuya, M.  
Mol. Cell. Biol. 6, 3000-3004, 1986  
A;Title: Human c-ros-1 gene homologous to the v-ros sequence of UR2 sarcoma virus encod  
A;Reference number: A25223; MUID:87064611; PMID:3023956  
A;Accession: A25223  
A;Molecule type: DNA  
A;Residues: 1790-2245, 'KFDGSEPSFRCTVN' <MA2>  
A;Cross-references: GB:M13368  
A;Experimental source: placenta  
A;Note: the differences after residue 2245 result from the authors' misinterpretation of  
R;Birchmeier, C.; Birnbaum, D.; Waitches, G.; Fasano, O.; Wigler, M.  
Mol. Cell. Biol. 6, 3109-3116, 1986  
A;Title: Characterization of an activated human ros gene.  
A;Reference number: A24421; MUID:87064625; PMID:3785223  
A;Accession: A24421  
A;Molecule type: mRNA  
A;Residues: 1854-2261, 'A', 2263-2347 <BI2>  
A;Cross-references: GB:M13880; NID:g337482; PIDN:AAA36580.1; PID:g337483  
A;Experimental source: tumor cells  
A;Note: the mcf3 oncogene was formed by DNA rearrangement involving fusion of at least t

C:Genetics:  
A:Gene: GDB:ROS1  
A:Cross-references: GDB:120351; OMIM:165020  
A:Map position: 6q22-6q22  
C:Superfamily: Kinase-related protein tos; LDL receptor YWTD-containing repeat homology;  
C:Keywords: ATP; autophosphorylation; glycoprotein; kinase-related transforming protein;  
F:1-36/Domain: signal sequence #status predicted <SIG>  
F:37-2347/Product: kinase-related protein ROS1 #status predicted <MAT>  
F:37-1859/Domain: extracellular #status predicted <EXT>  
F:335-378/Domain: LDL receptor YWTD-containing repeat homology <YW1>  
F:466-503/Domain: LDL receptor YWTD-containing repeat homology <YWA>  
F:715-757/Domain: LDL receptor YWTD-containing repeat homology <YW2>  
F:758-798/Domain: LDL receptor YWTD-containing repeat homology <YW3>  
F:799-838/Domain: LDL receptor YWTD-containing repeat homology <YW4>  
F:843-888/Domain: LDL receptor YWTD-containing repeat homology <YW5>  
F:893-933/Domain: LDL receptor YWTD-containing repeat homology <YW6>  
F:1532-1574/Domain: LDL receptor YWTD-containing repeat homology <YW7>  
F:1860-1883/Domain: transmembrane #status predicted <TMN>  
F:1884-2347/Domain: intracellular #status predicted <INT>  
F:1943-2222/Domain: protein kinase homology <KIN>  
F:1951-1959/Region: protein kinase ATP-binding motif  
F:52,114,123,324,352,471,607,628,706,714,732,939,961,1015,1087,1090,1211,1272,1330,1458,  
F:1980/Active site: Lys #status predicted  
F:2110,2114,2115/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation) #stat

Alignment Scores:  
Pred. No.: 8.98 Length: 2347  
Score: 9.00 Matches: 9  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 0.80% Indels: 0  
DB: 1 Gaps: 0

US-09-270-437D-6 (1-3412) x TVHURS (1-2347)

QY 569 ACCGAGCCAGCGTGGGACCACTCTT 595  
Db 661 SerGluProSerValGlyThrThrLeu 669  
|||||  
|||||

RESULT 14  
T12121  
NADH dehydrogenase 4 - Atlantic horseshoe crab mitochondrion (fragment)  
C:Species: mitochondrion Limulus polyphemus (Atlantic horseshoe crab)  
C:Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 11-May-2000  
C:Accession: T12121  
R:Staton, J.L.; Daehler, L.L.; Brown, W.M.  
Mol. Biol. Evol. 14, 867-874, 1997  
A:Title: Mitochondrial gene arrangement of the horseshoe crab Limulus polyphemus L.: Con  
A:Reference number: Z17427; MUID:97398711; PMID:9254925  
A:Accession: T12121  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-63 <STA>  
A:Cross-references: EMBL:AF002647; NID:g2316048; PID:g2316055; PIDN:AAC47689.1  
C:Genetics:  
A:Genome: mitochondrion  
A:Note: ND4  
C:Keywords: mitochondrion

Alignment Scores:  
Pred. No.: 144 Length: 63  
Score: 8.00 Matches: 8  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 0.73% Indels: 0  
DB: 2 Gaps: 0

US-09-270-437D-6 (1-3412) x T12121 (1-63)

QY 1833 TTTCGTCTCTGCTGCTTCACCTG 1810  
Db 39 PheLeuLeuLeuLeuHisLeu 46  
|||||  
|||||

RESULT 15  
F72580  
hypothetical protein APE1925 - Aeropyrum pernix (strain K1)  
C:Species: Aeropyrum pernix  
C:Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 20-Aug-1999  
C:Accession: F72580  
R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Taka  
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.;  
DNA Res. 6, 83-101, 1999  
A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropy  
A:Reference number: A72450; MUID:99310339; PMID:10382966  
A:Accession: F72580  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-121 <KAW>  
A:Cross-references: DDBJ:AF000062; NID:g5105244; PIDN:BAA80931.1; PID:d1044717; PID:g51  
A:Experimental source: strain K1  
C:Genetics:  
A:Gene: APE1925

Alignment Scores:  
Pred. No.: 132 Length: 121  
Score: 8.00 Matches: 8  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 0.73% Indels: 0  
DB: 2 Gaps: 0

US-09-270-437D-6 (1-3412) x F72580 (1-121)

QY 2058 CTTCCCGCCCTCCTCGGCCCT 2035  
Db 95 LeuProArgProSerSerAlaPro 102  
|||||  
|||||

Search completed: July 16, 2004, 11:32:56  
Job time : 82.5 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: July 16, 2004, 10:45:36 ; Search time 32 Seconds  
(without alignments)  
11103.961 Million cell updates/sec

Title: US-09-270-437D-6

Perfect score: 1120

Sequence: 1 ggcaggaggaggcgaggga.....aaccttgaaatgtttattt 3412

Scoring table: OLIGO

Xgapop 60.0 , Xgapext 60.0  
Ygapop 60.0 , Ygapext 60.0  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 141681 seqs, 52070155 residues

Word size: 1

Total number of hits satisfying chosen parameters: 283186

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODE=frame+\_n2p.model -DRV=xlp  
-Q=/cgn2\_1/USPTO.spool\_p/US09270437/rumat\_16072004\_113127\_13943/app.query.fasta\_1.3591  
-DB=SwissProt 42 -QMT=fastan -SUFFIX=Oligo.rsp -MINMATCH=0.1 -LOOPCL=0  
-LOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi  
-LIST=45 -DOALIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09270437.acgn\_1\_1\_46\_rumat\_16072004\_113127\_13943 -NCPUS=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGIOG  
-DRV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	9	0.8	398	1 KR2_PRVN3	P30662 pseudorabie
2	9	0.8	491	1 FIBR_HUMAN	P02675 homo sapien
3	9	0.8	781	1 SPI_MOUSE	O89090 mus musculu
4	9	0.8	785	1 SPI_HUMAN	P08047 homo sapien
5	9	0.8	788	1 SPI_RAT	Q01714 rattus norv
6	9	0.8	2347	1 KROS_HUMAN	P08922 homo sapien
7	8	0.7	111	1 FTSE_RALSO	O8Y0B4 ralstonia s
8	8	0.7	137	1 N551_SOYBN	Q05544 glycine max
9	8	0.7	150	1 SHIB_PIG	P79399 sus scrofa
10	8	0.7	161	1 SHIB_CANFA	P79250 canis fami
11	8	0.7	171	1 Y175_STAAM	O53719 staphylococ
12	8	0.7	202	1 R77_NEUCR	O43105 neurospora
13	8	0.7	208	1 YSX2_CAEEL	Q10021 caenorhabdi
14	8	0.7	224	1 RS2_METEA	O8PW41 methanosarc
15	8	0.7	225	1 CSMI_METAC	O8tt39 methanosarc
16	8	0.7	244	1 CSMI_CHLTE	O68988 chlorobium
17	8	0.7	285	1 Y891_MYCTU	Q10551 mycobacteri
18	8	0.7	285	1 Y915_MYCBO	P59970 mycobacteri

C 19	8	0.7	303	1 GDFF_MOUSE	Q92017 mus musculu
C 20	8	0.7	303	1 GDFF_RAT	Q92016 rattus norv
C 21	8	0.7	311	1 RBFI_MOUSE	P22560 mus musculu
C 22	8	0.7	316	1 TAL_HELPJ	O92jc5 helicobacte
C 23	8	0.7	326	1 YJHS_ECOLI	P39370 escherichia
C 24	8	0.7	379	1 CYB_PENFU	Q9blx4 pentagalagus
C 25	8	0.7	381	1 CYB_NOTTY	O03478 notoryctes
C 26	8	0.7	386	1 SHIB_CRIGR	P46636 cricetus
C 27	8	0.7	386	1 SHIB_MOUSE	P28334 mus musculu
C 28	8	0.7	386	1 SHIB_RAT	P28564 rattus norv
C 29	8	0.7	386	1 SHIB_SPAEH	P56496 spalax leuc
C 30	8	0.7	388	1 SHIB_DIDMA	P35404 didelphis m
C 31	8	0.7	389	1 SHIB_CAVPO	O08892 cavia porce
C 32	8	0.7	390	1 SHIB_HUMAN	P28222 homo sapien
C 33	8	0.7	390	1 SHIB_PANTR	P60020 pan troglod
C 34	8	0.7	390	1 SHIB_RABIT	P49144 oryctolagus
C 35	8	0.7	395	1 UMPI_ARATH	O9lka5 arabidopsis
C 36	8	0.7	407	1 IE68_HSVSA	Q01042 herpesvirus
C 37	8	0.7	411	1 IHH_MOUSE	P77812 mus musculu
C 38	8	0.7	414	1 SX17_HUMAN	Q9hg12 homo sapien
C 39	8	0.7	419	1 SX17_MOUSE	O61473 mus musculu
C 40	8	0.7	428	1 FXB2_MOUSE	O64733 mus musculu
C 41	8	0.7	428	1 HEMY_HAEIN	P44772 haemophilus
C 42	8	0.7	480	1 PRCP_HUMAN	P10619 homo sapien
C 43	8	0.7	498	1 VE2_HPVO8	P06422 human papil
C 44	8	0.7	499	1 PITE_ECOLI	P43676 escherichia
C 45	8	0.7	500	1 CP46_HUMAN	Q9y6a2 homo sapien

#### ALIGNMENTS

RESULT 1

KR2_PRVN3	KR2_PRVN3	STANDARD;	PRT;	398 AA.
AC	P30662;			
DT	01-APR-1993 (rel. 25, Created)			
DT	01-APR-1993 (rel. 25, Last sequence update)			
DT	01-OCT-1996 (rel. 34, Last annotation update)			
DE	Serine/threonine-protein kinase 2 (EC 2.7.1.-).			
GN	UL13 OR ULPK.			
OS	Pseudorabies virus (strain NIA-3) (PRV).			
OC	Viruses; dsDNA viruses, no RNA stage; Herpesviridae;			
OC	Alphaherpesvirinae; Varicellovirus.			
OX	NCBI_TaxID=10349;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=92365105; PubMed=1323689;			
RA	de Wind N., Domen J., Berns A.;			
RT	"Herpesviruses encode an unusual protein-serine/threonine kinase			
RT	which is nonessential for growth in cultured cells.";			
RL	J. Virol. 66:5200-5209(1992).			
CC	-!- FUNCTION: PROTEIN-SERINE/THREONINE KINASE NONESSENTIAL FOR GROWTH			
CC	IN CULTURED CELLS.			
CC	-!- PTM: Autophosphorylated (possible).			
CC				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC				
DR	EMBL; M94870; AAA47481.1; -.			
DR	PIR; B42744; WZBEN3.			
DR	InterPro; IPR000719; Prot kinase.			
DR	InterPro; IPR008271; Ser_thr_pkin_AS.			
DR	PROSITE; PS00107; PROTEIN KINASE ATP; FALSE_NEG.			
DR	PROSITE; PS00108; PROTEIN KINASE ST; 1.			
DR	PROSITE; PS50011; PROTEIN KINASE_DOM; 1.			
KW	Transferase; Serine/threonine-protein kinase; ATP-binding;			
KW	Phosphorylation; Manganese.			
FT	DOMAIN 4 9 POLY-GLY.			

FT DOMAIN 80 398 PROTEIN KINASE.  
 FT NP BIND 86 94 ATP (BY SIMILARITY).  
 FT BINDING 103 103 ATP (BY SIMILARITY).  
 FT ACT SITE 194 194 BY SIMILARITY.  
 SQ SEQUENCE 398 AA; 41416 MW; 45639AD75E42309B CRC64;

Alignment Scores:  
 Pred. No.: 4.87 Length: 398  
 Score: 9.00 Matches: 9  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 0.80% Indels: 0  
 DB: 1 Gaps: 0

US-09-270-437D-6 (1-3412) x KR2\_PRVN3 (1-398)

Qy 2 GCAGCGGAGGAGCGAGGAGCGCGGG 28

Db 119 AlaxlaGluGluAlaArgSerAlaGly 127

# RESULT 2

ID FIBB\_HUMAN STANDARD; PRT; 491 AA.  
 AC P02675;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 01-JUL-1993 (Rel. 26, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Fibrinogen beta chain precursor [Contains: Fibrinopeptide B].  
 GN FGB.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=91344740; PubMed=2102623;  
 RA Chung D.W., Harris J.E., Davie E.W.;  
 RT "Nucleotide sequences of the three genes coding for human  
 fibrinogen";  
 RL Adv. Exp. Med. Biol. 281:39-48(1990).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=83283433; PubMed=6688356;  
 RA Chung D.W., Que B.G., Rixon M.W., Mace M. Jr., Davie E.W.;  
 RT "Characterization of complementary deoxyribonucleic acid and genomic  
 deoxyribonucleic acid for the beta chain of human fibrinogen";  
 RL Biochemistry 22:3244-3250(1983).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA Chung D.W., Harris J.E., Davie E.W.;  
 RT "Nucleotide sequences of the three genes coding for human  
 fibrinogen";  
 RL (in) Liu C.Y., Chien S. (eds.);  
 RL Fibrinogen, thrombosis, coagulation and fibrinolysis, pp.39-48,  
 RL Plenum Press, New York (1991).  
 RN [4]  
 RP SEQUENCE FROM N.A., AND VARIANTS SER-100; HIS-170; LEU-265 AND  
 RP LYS-478.  
 RA Rieder M.J., Carrington D.P., Chung M.-W., Lee K.L., Yi Q.,  
 RA Nickerson D.A.;  
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.  
 RN [5]  
 RP SEQUENCE OF 31-491, AND CARBOHYDRATE-LINKAGE SITE.  
 RA Henschen A., Lottspeich F., Southan C., Topfer-Petersen E.;  
 RT "Human fibrinogen: sequence, sulfur bridges, glycosylation and some  
 structural variants";  
 RL (in) Peeters H. (eds.);  
 RL Protides of the biological fluids, Proc. 28th colloquium, pp.51-56,  
 RL Pergamon Press, Oxford (1980).  
 RN [6]  
 RP SEQUENCE OF 31-491.  
 RX MEDLINE=79124640; PubMed=420779;  
 RA Watt K.W.K., Takagi T., Doolittle R.F.;

RT "Amino acid sequence of the beta chain of human fibrinogen.";  
 RL Biochemistry 18:68-76(1979).  
 RN [7]  
 RP SEQUENCE OF 31-148, AND DISULFIDE BONDS.  
 RX MEDLINE=76225080; PubMed=936108;  
 RA Blomback B., Hessel B., Hogg D.;  
 RT "Disulfide bridges in NH2-terminal part of human fibrinogen.";  
 RL Thromb. Res. 8:639-658(1976).  
 RN [8]  
 RP SEQUENCE OF 1-38 FROM N.A.  
 RX MEDLINE=87146483; PubMed=3029722;  
 RA Huber P., Dalmon J., Courtois G., Laurent M., Assouline Z.,  
 RA Marguerie G.;  
 RT "Characterization of the 5'-flanking region for the human fibrinogen  
 beta gene";  
 RL Nucleic Acids Res. 15:1615-1625(1987).  
 RN [9]  
 RP SEQUENCE OF 31-44.  
 RA Blomback B., Blomback M., Grondahl N.J., Guthrie C., Hinton M.;  
 RT "Studies on fibrinopeptides from primates";  
 RL Acta Chem. Scand. 19:1788-1789(1965).  
 RN [10]  
 RP REVIEW, AND DISULFIDE BONDS.  
 RX MEDLINE=83254370; PubMed=6575689;  
 RA Henschen A., Lottspeich F., Kehl M., Southan C.;  
 RT "Covalent structure of fibrinogen";  
 RL Ann. N.Y. Acad. Sci. 408:28-43(1983).  
 RN [11]  
 RP DISULFIDE BONDS.  
 RX MEDLINE=77245999; PubMed=891553;  
 RA Gaardlund B., Hessel B., Marguerie G., Murano G., Blomback B.;  
 RT "Primary structure of human fibrinogen. Characterization of  
 disulfide-containing cyanogen-bromide fragments";  
 RL Eur. J. Biochem. 77:595-610(1977).  
 RN [12]  
 RP DISULFIDE BONDS.  
 RL Doolittle R.F., Takagi T., Watt K.W.K., Bouma H. III, Cottrell B.A.,  
 RA Cassman K.G., Goldbaum D.M., Doolittle L.R., Friesner S.J.;  
 RT "The structures of fibrinogen and fibrin";  
 RL (in) Magnusson S., Ottesen M., Foltmann B., Dano K.,  
 RL Neurath H. (eds.);  
 RL Regulatory proteolytic enzymes and their inhibitors, pp.163-172,  
 RL Pergamon Press, New York (1978).  
 RN [13]  
 RP REVIEW, EM STRUCTURE, POLYMERIZATION, AND LIGANDS.  
 RX MEDLINE=84305751; PubMed=6383194;  
 RA Doolittle R.F.;  
 RT "Fibrinogen and fibrin";  
 RL Annu. Rev. Biochem. 53:195-229(1984).  
 RN [14]  
 RP X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS) OF 164-491.  
 RX MEDLINE=97472408; PubMed=933233;  
 RA Spraggon G., Everse S.J., Doolittle R.F.;  
 RT "Crystal structures of fragment D from human fibrinogen and its  
 crosslinked counterpart from fibrin";  
 RL Nature 389:455-462(1997).  
 RN [15]  
 RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF 164-491.  
 RX MEDLINE=98292395; PubMed=9628725;  
 RA Everse S.J., Spraggon G., Veerapandian L., Riley M., Doolittle R.F.;  
 RT "Crystal structure of fragment Double-D from human fibrin with two  
 different bound ligands";  
 RL Biochemistry 37:8637-8642(1998).  
 RN [16]  
 RP X-RAY CRYSTALLOGRAPHY.  
 RX MEDLINE=99175089; PubMed=10074346;  
 RA Everse S.J., Spraggon G., Veerapandian L., Doolittle R.F.;  
 RT "Conformational changes in fragments D and double-D from human  
 fibrin(ogen) upon binding the peptide ligand Gly-His-Arg-Pro-amide";  
 RL Biochemistry 38:2941-2946(1999).  
 RN [17]  
 RP INTERACTION WITH FBLN1.  
 RX MEDLINE=95370284; PubMed=7642629;



RA Tran H., Tanaka A., Litvinovich S.V., Medved L.V., Haudenschild C.C.,  
 RA Agraves W.S.;  
 RT "The interaction of fibulin-1 with fibrinogen. A potential role in  
 RT hemostasis and thrombosis.";  
 RL J. Biol. Chem. 270:19458-19464(1995).  
 [18]  
 RN VARIANT BALTIMORE-2 LYS-478.  
 RP MEDLINE=89058942; PubMed=3194892;  
 RA Schmelzer C.H., Ebert R.F., Bell W.R.;  
 RT "A polymorphism at B beta 448 of fibrinogen identified during  
 RT structural studies of fibrinogen Baltimore II.";  
 RL Thromb. Res. 52:173-177(1988).  
 [19]  
 RN VARIANT ISE ARG-45.  
 RP MEDLINE=91208409; PubMed=2018836;  
 RA Yoshida N., Wada H., Morita K., Hirata H., Matsuda M., Yamazumi K.,  
 RA Asakura S., Shirakawa S.;  
 RT "A new congenital abnormal fibrinogen Ise characterized by the  
 RT replacement of B beta glycine-15 by cysteine.";  
 RL Blood 77:1958-1963(1991).  
 [20]  
 RN VARIANT NAPLES THR-98.  
 RP MEDLINE=92340664; PubMed=1634610;  
 RA Koopman J., Haverkate F., Lord S.T., Grimbergen J., Mannucci P.M.;  
 RT "Molecular basis of fibrinogen Naples associated with defective  
 RT thrombin binding and thrombophilia. Homozygous substitution of B beta  
 RT 68 Ala-->Thr.";  
 RL J. Clin. Invest. 90:238-244(1992).  
 [21]  
 RN VARIANTS IJMUUDEN CYS-44 AND NIJMEGEN CYS-74.  
 RP MEDLINE=92228809; PubMed=1565641;  
 RA Koopman J., Haverkate F., Grimbergen J., Engesser L., Novakova I.,  
 RA Kerst A.F.J.A., Lord S.T.;  
 RT "Abnormal fibrinogens IJmuuden (B beta Arg14-->Cys) and Nijmegen (B  
 RT beta Arg44-->Cys) form disulfide-linked fibrinogen-albumin  
 RT complexes.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 89:3478-3482(1992).  
 [22]  
 RN VARIANT NEW YORK-1 39-GLY--LEU-102 DEL.  
 RP MEDLINE=85157605; PubMed=3156856;  
 RA Liu C.Y., Koehn J.A., Morgan F.J.;  
 RT "Characterization of fibrinogen New York 1. A dysfunctional  
 RT fibrinogen with a deletion of B beta(9-72) corresponding exactly to  
 RT exon 2 of the gene.";  
 RL J. Biol. Chem. 260:4390-4396(1985).  
 [23]  
 RN VARIANTS GLU-2; LEU-265 AND LYS-478.  
 RP MEDLINE=99318093; PubMed=10391209;  
 RA Cargill M., Altschuler D., Ireland J., Sklar P., Ardlie K., Patil N.,  
 RA Shaw N., Lane C.R., Lim E.P., Kalyanaraman N., Nemesh J., Ziaugra L.,  
 RA Friedland L., Rolfe A., Warrington J., Lipshutz R., Daley G.Q.,  
 RA Lander E.S.;  
 RT "Characterization of single-nucleotide polymorphisms in coding regions  
 RT of human genes.";  
 RL Nat. Genet. 22:231-238(1999).  
 [24]  
 RN ERRATUM.  
 RP Cargill M., Altschuler D., Ireland J., Sklar P., Ardlie K., Patil N.,  
 RA Shaw N., Lane C.R., Lim E.P., Kalyanaraman N., Nemesh J., Ziaugra L.,  
 RA Friedland L., Rolfe A., Warrington J., Lipshutz R., Daley G.Q.,  
 RA Lander E.S.;  
 RL Nat. Genet. 23:373-373(1999).  
 [25]  
 RN VARIANTS CONGENITAL AFIBRINOGENEMIA ARG-383 AND ASP-430.  
 RP MEDLINE=20129589; PubMed=10666208;  
 RA Duga S., Asselta R., Santagostino E., Zeinali S., Simonic T.,  
 RA Malcovati M., Mannucci P.M., Turchini M.L.;  
 RT "Missense mutations in the human beta fibrinogen gene cause  
 RT congenital afibrinogenemia by impairing fibrinogen secretion.";  
 RL Blood 95:1336-1341(2000).  
 [26]  
 RN VARIANT CONGENITAL AFIBRINOGENEMIA CYS-196.  
 RP MEDLINE=21361164; PubMed=11468164;

RA Lounes K.C., Lefkowitz J.B., Henschen-Edman A.H., Coates A.I.,  
 RA Hantgan R.R., Lord S.T.;  
 RT "The impaired polymerization of fibrinogen Longmont  
 Alignment Scores:  
 Pred. No.: 4.79 Length: 491  
 Score: 9.00 Matches: 9  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 0.82% Indels: 0  
 DB: 1 Gaps: 0  
 US-09-270-437D-6 (1-3412) x FIBB\_HUMAN (1-491)  
 QY 3171 CTTCTCTGTGTCTTTTCTCGTTAAA 3145  
 |||||  
 Db 21 LeuLeuLeuCysValPheLeuValLys 29  
 RESULT 3  
 SPI\_MOUSE  
 ID SPI\_MOUSE STANDARD; PRT; 781 AA.  
 AC O89090; Q62251; Q64167;  
 DT 10-OCT-2003 (Rel. 42, Created)  
 DT 10-OCT-2003 (Rel. 42, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Transcription factor Sp1.  
 GN SP1.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A. (ISOFORM 1).  
 RC TISSUE=Neuroblastoma;  
 RX MEDLINE=98290554; PubMed=9628590;  
 RA Vajina S., Lee S.H., Minowa T., Mouradian M.M.;  
 RT "Sp family transcription factors regulate expression of rat D2  
 RT dopamine receptor gene.";  
 RL DNA Cell Biol. 17:471-479(1998).  
 [2]  
 RN SEQUENCE FROM N.A. (ISOFORM 2).  
 RX MEDLINE=96016118; PubMed=7568082;  
 RA Persengiev S.P., Saffer J.D., Kilpatrick D.L.;  
 RT "An alternatively spliced form of the transcription factor Sp1  
 RT containing only a single glutamine-rich transactivation domain.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 92:9107-9111(1995).  
 [3]  
 RN SEQUENCE OF 681-781 FROM N.A.  
 RP MEDLINE=92338398; PubMed=1633330;  
 RA Chestier A., Charnay P.;  
 RT "Difference in the genomic organizations of the related transcription  
 RT factors Sp1 and Krox-20; possible evolutionary significance.";  
 RL DNA Seq. 2:325-327(1992).  
 CC -!- FUNCTION: BINDS TO GC BOX PROMOTERS ELEMENTS AND SELECTIVELY  
 CC ACTIVATES MRNA SYNTHESIS FROM GENES THAT CONTAIN FUNCTIONAL  
 CC RECOGNITION SITES.  
 CC -!- SUBCELLULAR LOCATION: Nuclear.  
 CC -!- ALTERNATIVE PRODUCTS:  
 CC Event=Alternative splicing; Named isoforms=2;  
 CC Name=1;  
 CC IsoId=O89090-1; Sequence=Displayed;  
 CC Name=2; Synonyms=Sp1-S;  
 CC IsoId=O89090-2; Sequence=VSP 007376;  
 CC -!- PTM: O-glycosylated; contains N-acetylglucosamine side chains (By  
 CC similarity).  
 CC -!- SIMILARITY: BELONGS TO THE SP1 FAMILY OF C2H2-TYPE ZINC-FINGER  
 CC PROTEINS.  
 CC -!- SIMILARITY: Contains 3 C2H2-type zinc fingers.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
 DR EMBL; AF062566; AAC16484.1; -.  
 DR EMBL; S79832; AAB35321.1; -.  
 DR EMBL; X60136; CAA42721.1; -.  
 DR HSPF; P08047; 1SP1.  
 DR MGO; MGI:98372; Spt.  
 DR GO; GO:0005634; C:nucleus; IDA.  
 DR GO; GO:0003700; F:transcription factor activity; IMP.  
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IMP.  
 DR InterPro; IPR007087; Znf\_C2H2.  
 DR Pfam; PF00096; zf-C2H2; 3.  
 DR ProDom; PD000003; Znf\_C2H2; 2.  
 DR SMART; SM00355; Znf\_C2H2; 3.  
 DR PROSITE; PS00028; ZINC\_FINGER\_C2H2\_1; 3.  
 DR PROSITE; PS0157; ZINC\_FINGER\_C2H2\_2; 3.  
 KW Transcription regulation; Activator; Zinc-finger; Metal-binding;  
 KW DNA-binding; Nuclear protein; Repeat; Glycoprotein;  
 KW Alternative splicing.  
 FT ZN\_FING 624 648 C2H2-TYPE 1.  
 FT ZN\_FING 654 678 C2H2-TYPE 2.  
 FT ZN\_FING 684 706 C2H2-TYPE 3.  
 FT VARSPLIC 57 370 Missing (in isoform 2).  
 FT /FTId=Vsp\_007376.  
 FT CONFLICT 459 459 V -> G (IN REF. 2).  
 SQ SEQUENCE 781 AA; 80486 MW; 14CD12B8C58CF921 CRC64;

Alignment Scores:  
 Pred. No.: 4,6 Length: 781  
 Score: 9,00 Matches: 9  
 Percent Similarity: 100,00% Conservative: 0  
 Best Local Similarity: 100,00% Mismatches: 0  
 Query Match: 0,82% Indels: 0  
 DB: 1 Gaps: 0

US-09-270-437D-6 (1-3412) x SPL\_MOUSE (1-781)  
 QY 893 ATCTCTTCGGCTAGTTTGTCATCT 867  
 DB 304 IleSerSerAlaSerLeuValSerSer 312

RESULT 4  
 ID SPL\_HUMAN STANDARD; PRT: 785 AA.  
 AC P08047; Q9H305; Q9NR51; Q9NY21; Q9NYE7;  
 DT 01-AUG-1988 (Rel. 08, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Transcription factor Spt.  
 GN SPL OR TSFPL.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE OF 4-785 FROM N.A.  
 RC TISSUE=Cervical carcinoma;  
 RA Haggart M.H., Ladurner A.G.;  
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE OF 1-558 FROM N.A.  
 RX MEDLINE=20545561; PubMed=10973950;  
 RA Takahara T., Kanazu S., Yanagisawa S., Akanuma H.;  
 RT "Heterogeneous Spt mRNAs in human HepG2 cells include a product of  
 RT homotypic trans-splicing."  
 RL J. Biol. Chem. 275:38067-38072 (2000).  
 RN [3]  
 RP SEQUENCE OF 90-785 FROM N.A., AND SEQUENCE OF 359-375 AND 670-675.  
 RX MEDLINE=88080466; PubMed=319186;  
 RA Kadosh J.T., Carner K.R., Masiaz F.R., Tjian R.;  
 RT "Isolation of cDNA encoding transcription factor Spt and functional

analysis of the DNA binding domain.";  
 RL Cell 51:1079-1090(1987).  
 RN [4]  
 RP SEQUENCE OF 1-109 FROM N.A.  
 RA Nicolas M., Noe V., Ciudad C.J.;  
 RT "Expression of transcription factor Spt mRNA in mammalian cells."  
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.  
 RN [5]  
 RP SEQUENCE OF 1-98 FROM N.A.  
 RA Handschug K., Huebner A.;  
 RT "Sequencing of the 5' end of human transcription factor Spt mRNA.";  
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.  
 RN [6]  
 RP O-GLYCOSYLATION.  
 RX MEDLINE=89003041; PubMed=3139301;  
 RA Jackson S.P., Tjian R.;  
 RT "O-glycosylation of eukaryotic transcription factors: implications  
 RT for mechanisms of transcriptional regulation.";  
 RL Cell 55:125-133(1988).  
 RN [7]  
 RP STRUCTURE BY NMR OF 654-684 AND 684-712.  
 RX MEDLINE=97218212; PubMed=9065444;  
 RA Narayan V.A., Kriwacki R.W., Cardonna J.P.;  
 RT "Structures of zinc finger domains from transcription factor Spt.  
 RT Insights into sequence-specific protein-DNA recognition.";  
 RL J. Biol. Chem. 272:7801-7809(1997).  
 RN [8]  
 RP IDENTIFICATION OF SEROTONIN 1A RECEPTOR PROMOTER BINDING SITES.  
 RX MEDLINE=96224025; PubMed=8626793;  
 RA Parks C.L., Shenk T.;  
 RT "The serotonin 1a receptor gene contains a TATA-less promoter that  
 RT responds to MAZ and Spt.";  
 RL J. Biol. Chem. 271:4417-4430(1996).  
 CC -!- FUNCTION: BINDS TO GC BOX PROMOTERS ELEMENTS AND SELECTIVELY  
 CC ACTIVATES MRNA SYNTHESIS FROM GENES THAT CONTAIN FUNCTIONAL  
 CC RECOGNITION SITES. CAN INTERACT WITH G/C-RICH MOTIFS FROM  
 CC SEROTONIN RECEPTOR PROMOTER.  
 CC -!- SUBCELLULAR LOCATION: Nuclear.  
 CC -!- PTM: O-glycosylated; contains N-acetylglucosamine side chains.  
 CC -!- SIMILARITY: BELONGS TO THE SPT FAMILY OF C2H2-TYPE ZINC-FINGER  
 CC PROTEINS.  
 CC -!- SIMILARITY: Contains 3 C2H2-type zinc fingers.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; AF252284; AAF67726.1; -.  
 EMBL; AB039286; BAB13476.1; -.  
 EMBL; J03133; AAA61154.1; -.  
 EMBL; AF255682; AAF78781.1; -.  
 EMBL; AJ272134; CAA75345.1; -.  
 PIR; A29635; A29635.  
 DR PDB; 1SP1; 21-APR-97.  
 DR PDB; 1SP2; 21-APR-97.  
 DR TRANSFAC; T00759; -.  
 DR GlycoSuiteDB; P08047; -.  
 DR Genew; HGNC:11205; SPT.  
 DR MIM; 189906; -.  
 DR GO; GO:0005634; C:nucleus; NAS.  
 DR GO; GO:0003702; F:RNA polymerase II transcription factor acti. . .; NAS.  
 DR GO; GO:0006355; P:transcriptional activator activity; NAS.  
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; NAS.  
 DR InterPro; IPR007087; Znf\_C2H2.  
 DR Pfam; PF00096; zf-C2H2; 3.  
 DR ProDom; PD000003; Znf\_C2H2; 2.  
 DR SMART; SM00355; Znf\_C2H2; 3.  
 DR PROSITE; PS00028; ZINC\_FINGER\_C2H2\_1; 3.  
 DR PROSITE; PS00028; ZINC\_FINGER\_C2H2\_2; 3.  
 DR PROSITE; PS00157; ZINC\_FINGER\_C2H2\_2; 3.

KW Transcription regulation; Activator; Zinc-finger; Metal-binding;  
KW DNA-binding; Nuclear protein; Repeat; Glycoprotein; 3D-structure.  
FT ZN FINGER 626 650 C2H2-TYPE 1.  
FT ZN FINGER 656 680 C2H2-TYPE 2.  
FT ZN FINGER 686 708 C2H2-TYPE 3.  
FT CONFLICT 366 366 D -> G (IN REF. 3; AA SEQUENCE).  
FT CONFLICT 670 670 S -> F (IN REF. 3; AA SEQUENCE).  
FT STRAND 657 657  
FT TURN 661 662  
FT STRAND 666 666  
FT HELIX 670 677  
FT TURN 678 680  
SQ SEQUENCE 785 AA; 80693 MW; 43893DBP6518B9EA CRC64;

Alignment Scores:  
Pred. No.: 4.6 Length: 785  
Score: 9.00 Matches: 9  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 0.82% Indels: 0  
DB: 1 Gaps: 0

US-09-270-437D-6 (1-3412) x SP1\_HUMAN (1-785)

QY 893 ATCTCTCGCTAGTTGGTCTCATCT 867  
Db 305 IIESSerAlaSerLeuValSerSer 313  
|||||

## RESULT 5

SP1\_RAT  
ID SP1\_RAT STANDARD; PRT; 788 AA.  
AC Q01714;  
DT 01-APR-1993 (Rel. 25, Created)  
DT 01-APR-1993 (Rel. 25, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Transcription factor Sp1.  
GN SP1.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP TISSUE=Liver;  
RC  
RX MEDLINE=93010958; PubMed=1356762;  
RA Imataka H., Sogawa K., Yasumoto K., Kikuchi Y., Sasano K.,  
Kobayashi A., Hayami M., Fujii-Kuriyama Y.;  
RT "Two regulatory proteins that bind to the basic transcription element  
(BTE), a GC box sequence in the promoter region of the rat P-4501A1  
gene";  
RL ENBO J. 11:3663-3671(1992).  
CC -!- FUNCTION: BINDS TO GC BOX PROMOTERS ELEMENTS AND SELECTIVELY  
CC ACTIVATES MRNA SYNTHESIS FROM GENES THAT CONTAIN FUNCTIONAL  
CC RECOGNITION SITES.  
CC -!- SUBCELLULAR LOCATION: Nuclear.  
CC -!- PTM: O-glycosylated; contains N-acetylglucosamine side chains (By  
CC similarity).  
CC -!- SIMILARITY: BELONGS TO THE SP1 FAMILY OF C2H2-TYPE ZINC-FINGER  
CC PROTEINS.  
CC -!- SIMILARITY: Contains 3 C2H2-type zinc fingers.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; D12768; BAA02235.1; -.  
CC PIR; JS0747; JS0747.  
CC HSSP; P08047; LSP1.  
CC TRANSFAC; T00754; -.

DR InterPro; IPR007087; Znf\_C2H2.  
DR Pfam; PF00096; zf-C2H2; 3.  
DR ProDom; PD000003; Znf\_C2H2; 2.  
DR SMART; SM00355; Znf\_C2H2; 3.  
DR PROSITE; PS00028; ZINC\_FINGER\_C2H2\_1; 3.  
DR PROSITE; PS0157; ZINC\_FINGER\_C2H2\_2; 3.  
KW Transcription regulation; Activator; Zinc-finger; Metal-binding;  
KW DNA-binding; Nuclear protein; Repeat; Glycoprotein.  
FT ZN FINGER 629 653 C2H2-TYPE 1.  
FT ZN FINGER 659 683 C2H2-TYPE 2.  
FT ZN FINGER 689 711 C2H2-TYPE 3.  
SQ SEQUENCE 788 AA; 81015 MW; AA2B0CAB81AAB80C CRC64;

Alignment Scores:  
Pred. No.: 4.6 Length: 788  
Score: 9.00 Matches: 9  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 0.82% Indels: 0  
DB: 1 Gaps: 0

US-09-270-437D-6 (1-3412) x SP1\_RAT (1-788)

QY 893 ATCTCTCGCTAGTTGGTCTCATCT 867  
Db 308 IIESSerAlaSerLeuValSerSer 316  
|||||

## RESULT 6

KROS\_HUMAN  
ID KROS\_HUMAN STANDARD; PRT; 2347 AA.  
AC P08922; Q15368;  
DT 01-NOV-1988 (Rel. 09, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Proto-oncogene tyrosine-protein kinase ROS precursor (BC 2.7.1.112)  
(c-ros-1).  
DE ROS1.  
GN ROS1.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=90280463; PubMed=2352949;  
RA Birchmeier C., O'Neill K., Riggs M., Wigler M.;  
RT "Characterization of ROS1 cDNA from a human glioblastoma cell line";  
RL Proc. Natl. Acad. Sci. U.S.A. 87:4799-4803(1990).  
RN [2]  
RP SEQUENCE OF 1790-2259 FROM N.A.  
RX MEDLINE=87064611; PubMed=3023956;  
RA Matsushime H., Wang L.-H., Shibuya M.;  
RT "Human c-ros-1 gene homologous to the v-ros sequence of UR2 sarcoma  
virus encodes for a transmembrane receptorlike molecule";  
RL Mol. Cell. Biol. 6:3000-3004(1986).  
RN [3]  
RP SEQUENCE OF 1854-2245 FROM N.A.  
RX MEDLINE=87064625; PubMed=3785223;  
RA Birchmeier C., Birnbaum D., Waitches G., Fasano O., Wigler M.;  
RT "Characterization of an activated human ros gene";  
RL Mol. Cell. Biol. 6:3109-3116(1986).  
CC -!- FUNCTION: THIS IS A PROBABLY A CELL GROWTH OR DIFFERENTIATION  
CC FACTOR RECEPTOR WITH A TYROSINE-PROTEIN KINASE ACTIVITY.  
CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein  
CC tyrosine phosphate.  
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
CC -!- SIMILARITY: Belongs to the Tyr family of protein kinases. Insulin  
CC receptor subfamily.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial



```
KW Cell division; Transmembrane; Inner membrane; Coiled coil;
KW Complete proteome.
FT DOMAIN 1 3 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 4 21 POTENTIAL.
FT DOMAIN 22 111 PERIPLASMIC (POTENTIAL).
FT DOMAIN 27 72 COILED COIL (POTENTIAL).
SQ SEQUENCE 111 AA; 12658 MW; E4A926F7359C6C9C CRC64;

Alignment Scores:
Pred. No.: 56.9 Length: 111
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.73% Indels: 0
DB: 1 Gaps: 0

US-09-270-437D-6 (1-3412) x FTSD_RALSO (1-111)
QY 475 CTATTTTGTCTCTCTCTCTGTTG 452
Db 6 LeuPheLeuLeuLeuLeuLeu 13

RESULT 8
N551_SOYBN STANDARD; PRT; 137 AA.
AC Q0554;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Early nodulin 55-1 precursor (N-55-1) (Fragment).
GN ENOD55-1.
OS Glycine max (Soybean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
OX NCBI_TaxID=3847;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Williams; TISSUE=Root;
RX MEDLINE=94003074; PubMed=9400132;
RA de Blank C., Mylona P., Katinakis P.C., Bisseling T., Franssen H.;
RT "Characterization of the soybean early nodulin cDNA clone GMEOD55.";
RL Plant Mol. Biol. 22:1167-1171(1993).
CC -1- SUBCELLULAR LOCATION: Peribacteroid membrane (Potential).
CC -1- DEVELOPMENTAL STAGE: Expressed at early stages of nodule
CC plants after which levels decrease.
CC -1- INDUCTION: During nodulation in legume roots after Rhizobium
CC infection, and after release of bacteria from the infection
CC thread.
CC -1- SIMILARITY: Contains 1 plastocyanin-like domain.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; X69156; CAA48908.1; -.
DR PIR; S37353; S37353.
DR InterPro; IPR008972; Cupredoxin.
DR InterPro; IPR003245; Plicyanin like.
DR Pfam; PF02298; Cu_bind_like; 1.
DR ProDom; PD003122; Plicyanin like; 1.
KW Nodulation; Glycoprotein; Nitrogen fixation; Signal; Membrane.
FT SIGNAL 1 ?
FT CHAIN <1 ? POTENTIAL.
FT DOMAIN ? 137 EARLY NODULIN 55-1.
FT PLASTOCYANIN-LIKE.
FT CARBOHYD 13 13 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 51 51 N-LINKED (GLCNAC. . .) (POTENTIAL).

KW CARBOHYD 68 68 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 137 AA; 14857 MW; 988DE6F89A2E0BE0 CRC64;

Alignment Scores:
Pred. No.: 55.9 Length: 137
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.71% Indels: 0
DB: 1 Gaps: 0

US-09-270-437D-6 (1-3412) x N551_SOYBN (1-137)
QY 2727 CCTCCCTCTCTCTCTCACTCCA 2750
Db 78 ProSerProSerProSerPro 85

RESULT 9
SH1B_PIG STANDARD; PRT; 150 AA.
AC P79399;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE 5-hydroxytryptamine 1B receptor (5-HT-1B) (Serotonin receptor) (5-HT1B) (Fragment).
GN HTR1B.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain cortex;
RA Wurch T., Lestienne F., Colpaert F.C., Pauwels P.J.;
RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: This is one of the several different receptors for 5-
CC hydroxytryptamine (serotonin), a biogenic hormone that functions
CC as a neurotransmitter, a hormone, and a mitogen. The activity of
CC this receptor is mediated by G proteins that inhibit adenylate
CC cyclase activity (By similarity).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
CC Strongest to the other 5HT-1 subtype receptors.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; Y11867; CAA72615.1; -.
DR InterPro; IPR000276; GPCR_Rhodopsn.
DR Pfam; PF00001; 7tm1; 1.
DR PRINTS; PR00237; GPCRHHODOPSN.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; FALSE_NEG.
DR PROSITE; PS00262; G_PROTEIN_RECEP_F1_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein;
FT NON_TER 1 1
FT DOMAIN <1 83 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 84 107 6 (POTENTIAL).
FT DOMAIN 108 116 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 117 141 7 (POTENTIAL).
FT DOMAIN 142 >150 CYTOPLASMIC (POTENTIAL).
FT NON_TER 150 150
SQ SEQUENCE 150 AA; 16941 MW; 084EDF34A349555A CRC64;

Alignment Scores:
Pred. No.: 55.4 Length: 150
Score: 8.00 Matches: 8
```

Percent Similarity: 100.00%      Conservative: 0  
Best Local Similarity: 100.00%      Mismatches: 0  
Query Match: 0.71%                Indels: 0  
DB: 1                                Gaps: 0

US-09-270-437D-6 (1-3412) x 5H1B\_PIG (1-150)

QY 2547 TCTGGATCACCTGTGTATGTCAAC 2570  
|||||

DB 49 SerGlySerProValTyValAsn 56

RESULT 10

SH1B CANFA

ID SH1B CANFA STANDARD; PRT; 161 AA.

AC P79250;

DT 15-JUL-1998 (Rel. 36, Created)

DT 15-JUL-1998 (Rel. 36, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE 5-hydroxytryptamine 1B receptor (5-HT-1B) (Serotonin receptor) (5-

HT1B) (5-HT1D subtype beta) (Fragment).

GN HT1B.

OS Canis familiaris (Dog).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.

OX NCBI\_TaxID=9615;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Beagle, and Alsatian; TISSUE=Artery;

RX MEDLINE=96296365; PubMed=8763409;

RA Sgard F., Faure C., Graham D.;

RT "Evidence for 5-HT1D beta but not 5-HT1D alpha receptor subtype

expression in canine large coronary arteries and saphenous vein.;"

RL Cardiovasc. Res. 31:793-799(1996)

CC -!- FUNCTION: This is one of the several different receptors for 5-

hydroxytryptamine (serotonin), a biogenic hormone that functions

as a neurotransmitter, a hormone, and a mitogen. The activity of

this receptor is mediated by G proteins that inhibit adenylyate

cyclase activity.

CC -!- SUBCELLULAR LOCATION: Integral membrane protein.

CC -!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.

CC Strongest to the other 5HT-1 subtype receptors.

-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----

CC EMBL; S82461; AAB37488.2; -.

DR InterPro; IPR000276; GPCR\_Rhodpsn.

DR Pfam; PF00001; 7tm.1; 1.

DR PRINTS; PR00237; GPCRHHODPSN.

DR PROSITE; PS00237; G-PROTEIN RECP\_F1\_1; PARTIAL.

DR PROSITE; PS0262; G-PROTEIN RECP\_F1\_2; 1.

KW G-protein coupled receptor; Transmembrane; Glycoprotein;

Multigene family; Lipoprotein; Palmitate.

FT NON\_TER 1

FT DOMAIN <1 13 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 14 37 5 (POTENTIAL).

FT DOMAIN 38 123 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 124 147 6 (POTENTIAL).

FT DOMAIN 148 156 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 157 >161 7 (POTENTIAL).

FT NON\_TER 161

SQ SEQUENCE 161 AA; 17978 MW; 7BA978FA08A025B1 CRC64;

Alignment Scores:

Pred. No.:

Score: 55.1      Length: 161

Percent Similarity: 100.00%      Matches: 8

Best Local Similarity: 100.00%      Conservative: 0

Mismatches: 0

Query Match: 0.71%      Indels: 0  
DB: 1                        Gaps: 0

US-09-270-437D-6 (1-3412) x 5H1B\_CANFA (1-161)

QY 2547 TCTGGATCACCTGTGTATGTCAAC 2570  
|||||

DB 89 SerGlySerProValTyValAsn 96

RESULT 11

YI75 STAAM

ID YI75 STAAM STANDARD; PRT; 171 AA.

AC Q53719;

DT 15-JUL-1998 (Rel. 36, Created)

DT 15-JUL-1998 (Rel. 36, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Hypothetical protein SAV1875/SA1692/MW1815 (ORF1).

GN SAV1875 OR SA1692 OR MW1815.

OS Staphylococcus aureus (strain Mu50 / ATCC 700699),

OS Staphylococcus aureus (strain N315),

OS Staphylococcus aureus (strain MW2), and

OS Staphylococcus aureus.

OC Bacteria; Firmicutes; Bacillales; Staphylococcus.

OX NCBI\_TaxID=158878, 158879, 196620, 1280;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Mu50 / ATCC 700699, and N315;

RX MEDLINE=21311952; PubMed=11418146;

RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,

RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,

RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,

RA Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,

RA Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,

RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,

RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;

RT "Whole genome sequencing of methicillin-resistant Staphylococcus

aureus.;"

RL Lancet 357:1225-1240(2001).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=MW2;

RX MEDLINE=22040717; PubMed=12044378;

RA Baba T., Takeuchi F., Kuroda M., Yuzawa H., Aoki K.-I., Oguchi A.,

RA Nagai Y., Iwama N., Asano K., Naimi T., Kuroda H., Cui L.,

RA Yanamoto K., Hiramatsu K.;

RT "Genome and virulence determinants of high virulence community-

acquired MRSA.;"

RL Lancet 359:1819-1827(2002).

RN [3]

RP SEQUENCE FROM N.A.

RC STRAIN=NCTC 8325;

RX MEDLINE=94131293; PubMed=7905453;

RA Borchardt S.A., Babwah A.V., Jayaswal R.K.;

RT "Sequence analysis of the region downstream from a peptidoglycan

hydrolase-encoding gene from Staphylococcus aureus NCTC8325.;"

RL Gene 137:253-258(1993).

CC -!- SIMILARITY: BELONGS TO THE TH1J / PPPI FAMILY.

-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----

CC EMBL; AP003363; BAB58037.1; -.

DR EMBL; AP003335; BAB42961.1; -.

DR EMBL; AP004828; BAB95680.1; -.

DR EMBL; L19300; AAA18514.1; -.

DR PIR; B89975; B89975.

DR MEROPS; C56.UPW, -.

DR InterPro; IPR006286; Peptidase\_C56.

DR InterPro: IPR002818; ThiJ/Pfpi.  
DR Pfam: PF01965; DJ-1\_Pfpi; 1  
DR TIGRams: TIGR01382; Pfpi; 1.  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 171 AA; 18632 MW; A571A08B0FDD719 CRC64;

Alignment Scores:  
Pred. No.: 54.8 Length: 171  
Score: 8.00 Matches: 8  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 0.71% Indels: 0  
DB: 1 Gaps: 0

US-09-270-437D-6 (1-3412) x Y175\_STAAM (1-171)

QY 2527 GGAAGGACATCGGCGAGTCTG 2550  
|||||  
Db 119 GlyArgThrLeuThrAlaValLeu 126

## RESULT 12

ID RS7\_NEUCR STANDARD; PRT; 202 AA.  
AC O43105;  
DT 15-DEC-1998 (Rel. 37, Created)  
DT 15-DEC-1998 (Rel. 37, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE 40S ribosomal protein S7.  
GN RPS-7 OR CRPS-7.  
OS Neurospora crassa.  
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.  
OX NCBI\_TaxID=5141;  
RN [1]

SEQUENCE FROM N.A.

RC STRAIN=74-OR23-IVA;

RX MEDLINE=97174111; PubMed=9021131;

RA Vierula P.J.;

RT "Cloning and characterization of a Neurospora crassa ribosomal  
protein gene, crps-7.";

RL Curr. Genet. 31:139-143(1997).

CC -1- SIMILARITY: Belongs to the S7E family of ribosomal proteins.

-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----

EMBL: U73847; AAB94301.1; -.

DR PIR; T46586; T46586.

DR InterPro: IPR000554; Ribosomal\_S7E.

DR Pfam: PF01251; Ribosomal\_S7e; 1.

DR ProDom: PD006276; Ribosomal\_S7E; 1.

DR PROSITE: PS00948; RIBOSOMAL\_S7E; 1.

KW Ribosomal protein.

SQ SEQUENCE 202 AA; 22819 MW; 94DB756FFB37798 CRC64;

Alignment Scores:

Pred. No.: 54 Length: 202  
Score: 8.00 Matches: 8  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 0.71% Indels: 0  
DB: 1 Gaps: 0

US-09-270-437D-6 (1-3412) x RS7\_NEUCR (1-202)

QY 1817 AGAGGAGGAGCGAATACCCCTC 1840  
|||||

Db 111 SerSerArgSerArgAsnThrLeu 118

## RESULT 13

YSX2 CAEEL

ID YSX2 CAEEL STANDARD; PRT; 208 AA.

AC Q10021; Q95ZN0; Q95ZN1; Q95ZN2;

DT 01-OCT-1996 (Rel. 34, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Hypothetical RNA-binding protein T28D9.2 in chromosome II.

GN T28D9.2.

OS Caenorhabditis elegans.

OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;

OC Rhabditidae; Peloderinae; Caenorhabditis.

OX NCBI\_TaxID=6239;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Bristol N2;

RA Fulton L.;

RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.

RN REVISIONS, AND ALTERNATIVE SPLICING.

RA Waterston R.;

RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.

CC -1- ALTERNATIVE PRODUCTS:

CC Event-Alternative splicing; Named isoforms=4;

CC Comment=Experimental confirmation may be lacking for some  
CC isoforms;

CC Name=a;

CC IsoId=Q10021-1; Sequence=Displayed;

CC Name=b;

CC IsoId=Q10021-2; Sequence=VSP\_005908, VSP\_005910;

CC Name=c;

CC IsoId=Q10021-3; Sequence=VSP\_005906, VSP\_005907;

CC Name=d;

CC IsoId=Q10021-4; Sequence=VSP\_005909;

CC -1- SIMILARITY: Contains 1 RNA recognition motif (RRM) domain.

-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----

EMBL: U28738; AAG68314.2; -.

DR EMBL; U28738; AAG68399.2; -.

DR EMBL; U28738; AAK68400.1; -.

DR EMBL; U28738; AAK72066.2; -.

DR WormPep; T28D9.2a; CE29826.

DR WormPep; T28D9.2b; CE29827.

DR WormPep; T28D9.2c; CE28094.

DR WormPep; T28D9.2d; CE29828.

DR InterPro; IPR000504; RNA\_rec\_mot.

DR Pfam; PF00076; rrm; 1.

DR SMART; SM00360; RRM; 1.

DR PROSITE; PS0102; RRM; 1.

DR PROSITE; PS00030; RRM\_RNP\_1; FALSE NEG.

KW Hypothetical protein; RNA-binding; Alternative splicing.

FT DOMAIN 2 74

FT DOMAIN 84 205

FT VARSPPLIC 1 128

FT ASG/SER-RICH

FT MERLYLGKTPYNGHERDVERFLKGYKINNISMKYGFADF  
FT FEDSRDAEDACHLDGKTKWEGSMRLVEMARGKPGNDRH  
FT GSRSPRRSRSPRRSRTPRRSRDRSRDRSRSRSSSS  
FT RSRSP -> MLKTHAMIMKEKPKWALCALSKWLAENHAA  
FT MTVMDRDHDHDPALLDVVHLHQEDVPALVIANVLADPVP  
FT EAPDPDHQFVKVAZANPVRARKETLWNVQYISIHFSF  
FT SFHSYOKPLFRUKE (in isoform c).  
FT /FTId=VSP\_005906.

FT VARSPPLIC 129 147

FT Missing (in isoform c).  
FT /FTId=VSP\_005907.

FT VARSPPLIC 147 166

FT LKREASRSPLPAKDRSRT -> FHSSMRNQVHLQAHIAM  
FT AVT (in isoform b).



```
FT VARSPLIC 148 208 /FtId=VSP_005908.
FT KREASRSPLPAKDRSTRSGSPKNGGDRKRSVSRGRSH
FT SRDGSNRSVSPSPGSPKD -> QLRANISVSPSPDSD
FT SIKCKNYLIGSVKRRDVIHHCQKLEAFEDVLPORTAEIE
FT SAA (in isoform d).
FT /FtId=VSP_005909.
FT Missing (in isoform b).
FT /FtId=VSP_005910.
SQ SEQUENCE 208 AA; 23946 MW; A33CAF78D74A6F8 CRC64;

Alignment Scores:
Pred. No.: 53.9 Length: 208
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.73% Indels: 0
DB: 1 Gaps: 0

US-09-270-437d-6 (1-3412) x YXS2_CABEL (1-208)
QY 202 AGTCACGAGGCGTAGCGGACT 179
Db 93 SerProArgArgSerArgThr 100

RESULT 14
RS2_METMA STANDARD; PRT; 224 AA.
AC Q8P4L1;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 30S ribosomal protein S2P.
GN RPS2P OR MW1760.
OS Methanosarcina mazei (Methanosarcina frisia).
OC Archaea; Euryarchaeota; Methanomicrobia; Methanosarcinales;
OC Methanosarcinaceae; Methanosarcina.
OX NCBI_TaxID=2209;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Goel / Goli / ATCC BAA-199 / DSM 3647 / OCM 88;
RX MEDLINE=22120827; PubMed=12125824;
RA Deppenmeier U., Johann A., Hartsch T., Merkl R., Schmitz R.A.,
RA Martinez-Arias R., Henne A., Wierzer A., Baeumer S., Jacobi C.,
RA Brueggemann H., Lienard T., Christmann A., Boencke M., Steckel S.,
RA Bhattacharya A., Lykidis A., Overbeek R., Klenk H.-P., Gunsalus R.P.,
RA Fritz H.-J., Gottschalk G.
RT "The genome of Methanosarcina mazei: evidence for lateral gene
transfer between Bacteria and Archaea."
RL J. Mol. Microbiol. Biotechnol. 4:453-461(2002).
CC -!- SIMILARITY: Belongs to the S2P family of ribosomal proteins.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AE013412; AAM31456.1; -.
CC HAWAP; MF_00291; -.
CC InterPro; IPR001865; Ribosomal S2.
CC InterPro; IPR005707; Ribosomal S2_e/a.
CC Pfam; PF00318; Ribosomal S2_1.
CC PRINTS; PR00395; RIBOSOMALS2.
CC TIGRFAMs; TIGR01012; Sa_S2_E_A; 1.
CC PROSITE; PS00962; RIBOSOMAL_S2_1; 1.
CC PROSITE; PS00963; RIBOSOMAL_S2_2; 1.
CC KX Ribosomal protein; Complete proteome.
SQ SEQUENCE 224 AA; 24625 MW; 144D106450EA125E CRC64;

Alignment Scores:
Pred. No.: 53.6 Length: 224
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.73% Indels: 0
DB: 1 Gaps: 0
```

```
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.73% Indels: 0
DB: 1 Gaps: 0

US-09-270-437d-6 (1-3412) x RS2_METMA (1-224)
QY 2496 CGGGTGGCTCAAGAGCTCCTCTCT 2473
Db 83 ArgValAlaSerLysLeuLeuSer 90

RESULT 15
RS2_METAC STANDARD; PRT; 225 AA.
AC Q8T39;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 30S ribosomal protein S2P.
GN RPS2P OR MA0600.
OS Methanosarcina acetivorans.
OC Archaea; Euryarchaeota; Methanomicrobia; Methanosarcinales;
OC Methanosarcinaceae; Methanosarcina.
OX NCBI_TaxID=2214;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C2A / ATCC 35395 / DSM 2834;
RX MEDLINE=21929760; PubMed=11932238;
RA Galagan J.E., Nusbaum C., Roy A., Endrizzi M.G., Macdonald P.,
RA FitzHugh W., Calvo S., Engels R., Smirnov S., Atnoor D., Brown A.,
RA Allen N., Naylor J., Stange-Thomann N., DeArellano K., Johnson R.,
RA Linton L., McEwan P., McKernan K., Talamas J., Tirrell A., Ye W.,
RA Zimmer A., Barber R.D., Cann I., Graham D.E., Guss A.M.,
RA Hederich R., Ingram-Smith C., Kuetner H.C., Krzycki J.A.,
RA Leigh J.A., Li W., Liu J., Mukhopadhyay B., Reeve J.N., Smith K.,
RA Springer T.A., Umayam L.A., White O., White R.H., de Macario E.C.,
RA Ferry J.G., Jarrell K.F., Jing H., Macario A.J.L., Paulsen I.,
RA Pritchett M., Sowers K.R., Swanson R.V., Zinder S.H., Lander E.,
RA Metcalf W.W., Birren B.;
RT "The genome of Methanosarcina acetivorans reveals extensive metabolic
and physiological diversity."
RL Genome Res. 12:532-542(2002).
CC -!- SIMILARITY: Belongs to the S2P family of ribosomal proteins.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AE010721; AAM04044.1; ALT_INIT.
CC HAWAP; MF_00291; -.
CC InterPro; IPR001865; Ribosomal S2.
CC InterPro; IPR005707; Ribosomal S2_e/a.
CC Pfam; PF00318; Ribosomal S2_1.
CC PRINTS; PR00395; RIBOSOMALS2.
CC TIGRFAMs; TIGR01012; Sa_S2_E_A; 1.
CC PROSITE; PS00962; RIBOSOMAL_S2_1; 1.
CC PROSITE; PS00963; RIBOSOMAL_S2_2; 1.
CC KX Ribosomal protein; Complete proteome.
SQ SEQUENCE 225 AA; 24779 MW; 02C72D26A5F411F CRC64;

Alignment Scores:
Pred. No.: 53.5 Length: 225
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.73% Indels: 0
DB: 1 Gaps: 0
```

US-09-270-437D-6 (1-3412) x RS2\_METAC (1-225)

QY 2496 CGGGTGGCCTCAAAGCTCCTCTCT 2473

Db 84 ArgValAlaSerLysLeuLeuSer 91

Search completed: July 16, 2004, 11:18:44  
Job time : 50 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: July 16, 2004, 10:57:48 ; Search time 162.5 Seconds

(without alignments)  
13249.823 Million cell updates/sec

Title: US-09-270-437D-6

Perfect score: 1120

Sequence: 1 ggcagcgaggagcgagga.....aaccttgaaatgtttattt 3412

Scoring table:

OLIGO  
Xgapop 60.0 , Xgapext 60.0  
Ygapop 60.0 , Ygapext 60.0  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 1017041 seqs, 315518202 residues

Word size: 1

Total number of hits satisfying chosen parameters: 2033934

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlp  
-Q=/cgn2 1/USPTO\_model\_P/US09270437/runat 16072004 113127 13955/app query.fasta 1.3591  
-DB=SPTRMBL 25 -QFMT=fastan -SUFFIX=oligo -MNMATCH=0.1 -LOOPL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR\_SCORE=quality -THR\_MIN=1 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09270437 @CEN 1.1 283 @runat 16072004 113127 13955 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOF=6  
-FGAPEXT=7 -YGAPOF=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

SPTRMBL 25: \*  
1: sp\_archaea: \*  
2: sp\_bacteria: \*  
3: sp\_fungi: \*  
4: sp\_human: \*  
5: sp\_invertebrate: \*  
6: sp\_mammal: \*  
7: sp\_thc: \*  
8: sp\_organelle: \*  
9: sp\_phage: \*  
10: sp\_plant: \*  
11: sp\_rodent: \*  
12: sp\_virus: \*  
13: sp\_vertebrate: \*  
14: sp\_unclassified: \*  
15: sp\_rvirus: \*  
16: sp\_bacteriap: \*  
17: sp\_archaeap: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	357	31.9	556	4 Q9Y6M1	Q9y6m1 homo sapien

ID	Q9Y6M1	PRELIMINARY;	PRT;	556 AA.
AC	Q9Y6M1;			
DT	01-NOV-1999 (TREMBLrel. 12, Created)			
DT	01-NOV-1999 (TREMBLrel. 12, Last sequence update)			
DT	01-OCT-2003 (TREMBLrel. 25, Last annotation update)			
DE	Hepatocellular carcinoma autoantigen.			
GN	P62.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=99207072; PubMed=10190901;			
RA	Zhang J.Y., Chan E.K., Peng X.X., Tan E.M.;			
RT	"A novel cytoplasmic protein with RNA-binding motifs is an autoantigen			
RT	in human hepatocellular carcinoma.";			
RL	J. Exp. Med. 189:1101-1110(1999).			
CC	-!- SIMILARITY: CONTAINS 4 KH DOMAINS.			
DR	EMBL; AF057352; AAD31596.1; -			
DR	GO; GO:0005737; C:cytoplasm; TAS.			
DR	GO; GO:0003723; F:RNA binding; TAS.			
DR	GO; GO:0007345; P:embryogenesis and morphogenesis; TAS.			

#### ALIGNMENTS

RESULT 1

Q9Y6M1 ID Q9Y6M1 PRELIMINARY; PRT; 556 AA.

DR GO; GO:0009386; P:translational attenuation; TAS.

DR InterPro; IPR004087; KH dom.

DR InterPro; IPR004088; KH\_type\_1.

DR InterPro; IPR000504; RNA\_rec\_mot.

DR Pfam; PF00013; KH; 4.

DR Pfam; PF00076; xim; 2.

DR SMART; SM00322; KH; 4.

DR SMART; SM00360; RRM; 2.

DR PROSITE; PS50084; KH TYPE 1; 4.

DR PROSITE; PS50102; RRM; 2.

SQ SEQUENCE 556 AA; 61842 MW; 1EDEFB100443DDC4 CRC64;

Alignment Scores:

Pred. No.: 0 Length: 556  
Score: 357.00 Matches: 357  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 31.87% Indels: 0  
DB: 4 Gaps: 0

US-09-270-437D-6 (1-3412) x Q9Y6M1 (1-556)

QY 70 ATGATGACACAGCTTATACATCGGAACTGAGACCCCGCGGTACCGCGACGACCTCCGG 129

Db 1 MetMetAsnLysLeuTyrIleGlyAsnLeuSerProAlaValThrAlaAspLeuArg 20

QY 130 CAGCTCTTTGGGACAGGAAGTCGCCCTGGCGGACAGAGCTCTGAGTCCGGCTAC 189

Db 21 GlnLeuPheGlyAspArgLysLeuProLeuAlaGlyGlnValLeuLeuLysSerGlyTyr 40

QY 190 GCCTTCGTGGACTACCCCGACAGCAACTGGGCCATCCCGGCATCGAGACCCCTCTCGGGT 249

Db 41 AlaPheValAspTyrProAspGlnAsnTrpAlaIleArgAlaIleGluThrLeuSerGly 60

QY 250 AAAGTGGATTGCATGGGAAATCATGGAAGTTGATTAAGTCTCTAAAGACTAAGG 309

Db 61 LysValGluLeuHisGlyLysIleMetGluValAspTyrSerValSerLysLysLeuArg 80

QY 310 AGCAGGAAATTCAGATTCGAAACATCCCTCTCACCTGCAGTGGAGGTGTGATGGA 369

Db 81 SerArgLysIleGlnIleArgAsnIleProProHisLeuGlnTrpGluValLeuAspGly 100

QY 370 CTTTGGCTCAATATGGACAGTGGAGATCTGGAACAGTCAAGTCAACACACACAGAAC 429

Db 101 LeuLeuAlaGlnTyrGlyThrValGluAsnValGluGlnValAsnThrAspThrGluThr 120

QY 430 GCCCTGTCACGTCACATATGCACACAGACAGAGCAAGCAAAATAGCCATCGAGAAGCTA 489

Db 121 AlaValValAsnValThrTyrAlaThrArgGluGluAlaLysIleAlaMetGluLysLeu 140

QY 490 AGCGGGCATCAGTTTGAGAACTACTCCTTCAAGATTTCTACATCCCGGATGAAGAGTG 549

Db 141 SerGlyHisGlnPheGluAsnTyrSerPheLysIleSerTyrIleProAspGluVal 160

QY 550 AGCTCCCTTCGCCCTCAGCAGCCCGAGTGGGACCACTCTTCCCGGAGCAAGGC 609

Db 161 SerSerProSerProGlnArgAlaGlnArgGlyAspHisSerSerArgGluGlnGly 180

QY 610 CACGCCCTCGGGGACCTCTCAGCCAGACAGATGATTTCCCGTCGCGATCTGTCG 669

Db 181 HisAlaProGlyGlyThrSerGlnAlaArgGlnIleAspPheProLeuArgIleLeuVal 200

QY 670 CCCACCCAGTTTGTGTGGTCCCATCATCGAAAGGAGGCTTGACCAATAAGAAACATCACT 729

Db 201 ProThrGlnPheValGlyAlaIleIleGlyLysGluGlyLeuThrIleLysAsnIleThr 220

QY 730 AAGCAGACCCAGTCCCGGTAGATATCCATGAAAGAGAACTCTGAGCTGCAGAGAAG 789

Db 221 LysGlnThrGlnSerArgValAspIleHisArgLysGluAsnSerGlyAlaAlaGluLys 240

QY 790 CCTCTCACCATCCATGCCACCCAGAGGGGACTTCTGAAGCATGCCGCATGATCTTGAA 849

Db 241 ProValThrIleHisAlaThrProGluGlyThrSerGluAlaCysArgMetIleLeuGlu 260

QY 850 ATCATGCAGAAAGAGCGAGATGAGACCAAACTAGCCGAGAGATTCCTCTGAAAATCTTGG 909

Db 261 IleMetGlnLysGluAlaAspGluThrLysLeuAlaGluIleProLeuLysIleLeu 280

QY 910 GCACACAATGGCTTGGTGGAAAGACTGATTCGAAAGAGGCAGAAAATTTGAAGAAAATT 969

Db 281 AlahisasnGlyLeuValGlyArgleuIleGlyLysGluGlyArgasnLeuLysIle 300

QY 970 GAACATGAACACAGGACCAAGATACAAATCTCATCTTTGTCAGGATTTGAGCATATACAC 1029

Db 301 GluHisGluThrGlyThrLysIleThrIleSerSerLeuGlnAspLeuSerIleTyrAsn 320

QY 1030 CCGGAAGAACCATCAGTGTGAAGGACAGTTCAGGCTGTGCCAGTCTGAGATAGAG 1089

Db 321 ProGluArgThrIleThrValLysGlyThrValGluAlaCysAlaSerAlaGluIleGlu 340

QY 1090 ATTATGAAGAAAGCTGCGTGAGGCCCTTTGAAAATGATATGCTGGCTGTTTAAAC 1140

Db 341 IleMetLysLysLeuArgGluAlaPheGluAsnAspMetLeuAlaValAsn 357

RESULT 2

Q7TQF9

ID Q7TQF9 PRELIMINARY; PRT; 545 AA.

AC Q7TQF9;

DT 01-OCT-2003 (TrEMBLrel. 25, Created)

DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DE Hypothetical protein.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

OX NCBI\_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6; TISSUE=Brain;

RX MEDLINE=22388257; PubMed=12477932;

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,

RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.W., Schuler G.D.,

RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,

RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,

RA Brownstein M.J., Uesdin T.B., Toshiyuki S., Carninci P., Prange C.,

RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,

RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,

RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,

RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,

RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,

RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,

RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,

RA Krzywinski M.I., Skalski U., Smallos D.E., Schnerch A., Schein J.E.,

RA Jones S.J., Marra M.A.;

RT "Generation and initial analysis of more than 15,000 full-length human

and mouse cDNA sequences."

RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6; TISSUE=Brain;

RA Strausberg R.;

RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.

DR EMBL; BC054552; AAH54552.1; -.

KW Hypothetical protein.

SQ SEQUENCE 545 AA; 59679 MW; 88DD35FF30DF3091 CRC64;

Alignment Scores:

Pred. No.: 2,13e-82 Length: 545

Score: 88.00 Matches: 88

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 7.86% Indels: 0

DB: 11 Gaps: 0

```
US-09-270-437D-6 (1-3412) x Q7QP9 (1-545)
QY 628 TCTCAGCCAGACAGATTGATTTCCCGTCCGATCCTGTGTCACCCAGCCAGCTTTGTGCT 687
Db 179 SerGlnAlaArgGlnIleaspPheProLeuArgIleLeuValProThrGlnPheValGly 198
QY 688 GCATCATCGGAAAGAGCGCTTGACCAATAAGAACATCACTAAGCAGACCCAGTCCCGG 747
Db 199 AlaIleGlyLysGluGlyLeuThrIleLysAsnIleThrLysGlnThrGlnSerArg 218
QY 748 GTAGATATCATAGAAAGAGACTCTGGAGCTGCAGAGAGCTGCACCATCATGCC 807
Db 219 ValAspIleHisArgLysGluAsnSerGlyAlaAlaGluLysProValThrIleHisAla 238
QY 808 ACCCCAGAGGGGACTTCTGAAGCATGCCGATCATCTTTGAATCATGCAGAAAGAGGCA 867
Db 239 ThrProGluGlyThrSerGluAlaCysArgMetIleLeuGluIleMetGlnLysGluAla 258
QY 868 GATGACCAAACTAGCCGAGAG 891
Db 259 AspGluThrLysLeuAlaGluGlu 266
RESULT 3
Q9PW80 PRELIMINARY; PRT; 582 AA.
AC Q9PW80;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Vg1 RNA binding protein (Decapentaplegic and Vg-related 1, RNA binding
  protein).
GN DVRLRBP.
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RA Zhang Q., Yaniv K., Oberman F., Wolke U., Git A., Fromer M.,
RA Taylor W., Meyer D., Standart N., Raz E., Yisraeli J.K.;
RT "Vg1 RBP intracellular distribution and evolutionarily conserved
  expression suggest multiple roles during development.";
RL Mech. Dev. 0:0-0(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=AB; TISSUE=Body;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smalls D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra W.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
  and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=AB; TISSUE=Body;
RX STRAUSBERG R.;
```

Submitted (JAN-2003) to the EMBL/GenBank/DBDJ databases.  
-!- SIMILARITY: CONTAINS 4 KH DOMAINS.  
DR EMBL; AF161270; RAD45610.1; -.  
DR EMBL; BC045873; RAD45873.1; -.  
DR ZFIN; ZDB-GENE-000308-1; dvrlrbp.  
DR GO; GO:0003676; F:nucleic acid binding; IEA.  
DR InterPro; IPR004087; KH dom.  
DR InterPro; IPR004088; KH\_type\_1.  
DR InterPro; IPR00504; RNA\_rec\_mot.  
DR Pfam; PF00013; KH; 4.  
DR Pfam; PF00076; xrm; 2.  
DR SMART; SM00322; KH; 4.  
DR SMART; SM00360; RRM; 2.  
DR PROSITE; PS00864; KH\_TYPE\_1; 4.  
DR PROSITE; PS0102; RRM; 2.  
SQ SEQUENCE 582 AA; 63351 MW; 9DA663200681B306 CRC64;

Alignment Scores:  
Pred. No.: 3,76e-14 Length: 582  
Score: 23.00 Matches: 23  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 2.05% Indels: 0  
DB: 13 Gaps: 0

US-09-270-437D-6 (1-3412) x Q9PW80 (1-582)  
QY 1507 GGGCCACCGAAGCCAGCTTCAAGCCCGAGCGGATCTTTGGGAACTGGAAGAGGAA 1566  
Db 461 GlyProGluAlaGlnPheLysAlaGlnGlyArgIlePheGlyLysLeuLysGluGlu 480  
QY 1567 AACTTCTTT 1575  
Db 481 AsnPhePhe 483  
RESULT 4  
O57526 PRELIMINARY; PRT; 593 AA.  
ID O57526;  
AC O57526;  
DT 01-JUN-1998 (TrEMBLrel. 06, Created)  
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE KH domain-containing transcription factor B3.  
OS Xenopus laevis (African clawed frog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;  
OC Xenopodinae; Xenopus.  
OX NCBI\_TaxID=8355;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=92249652; PubMed=1577195;  
RA Pfaff S.L., Taylor W.L.;  
RT "Characterization of a Xenopus oocyte factor that binds to a  
 developmentally regulated cis-element in the TFIIA gene.";  
RL Dev. Biol. 151:306-316(1992).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Griffin D., Taylor W.L.;  
RT Submitted (JAN-1998) to the EMBL/GenBank/DBDJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RA Havin L., Git A., Elisha Z., Oberman F., Yaniv K.,  
RA Pressman Schwartz S., Standart N.M., Yisraeli J.K.;  
RL Genes Dev. 0:0-0(1998).  
CC -!- SIMILARITY: CONTAINS 4 KH DOMAINS.  
DR EMBL; AF042353; AAB97457.1; -.  
DR EMBL; AF064633; AAC18597.1; -.  
DR GO; GO:0003676; F:nucleic acid binding; IEA.  
DR InterPro; IPR004087; KH dom.  
DR InterPro; IPR004088; KH\_type\_1.  
DR InterPro; IPR00504; RNA\_rec\_mot.  
DR Pfam; PF00013; KH; 4.  
DR Pfam; PF00076; xrm; 2.

```

DR SMART; SM00322; KH; 4.
DR SMART; SM00360; RRM; 2.
DR PROSITE; PS00084; KH TYPE 1; 4.
DR PROSITE; PS0102; RRM; 2.
DR PROSITE; PS00030; RRM_RNP_1; FALSE NEG.
SQ SEQUENCE 593 AA; 65385 MW; 5A5AB4B4A1D55DF7 CRC64;

Alignment Scores:
Pred. No.: 3.75e-14 Length: 593
Score: 23.00 Matches: 23
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.05% Indels: 0
DB: 13 Gaps: 0

US-09-270-437D-6 (1-3412) x O57526 (1-593)

QY 1630 GCTGCGCGGTGATTGGCAAGAGTGGCAAGACCGTGAACGAACTCGAAGCTTAAACCACT 1689
Db |||||
QY 511 AlaGlyArgValIleGlyLysThrValAsnGluLeuGlnAsnLeuThrSer 530
Db |||||
QY 1690 GCAGAAAGTC 1698
Db |||||
QY 531 AlaGluVal 533
Db |||||

RESULT 5
O73932 PRELIMINARY; PRT; 594 AA.
AC O73932;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE VGI RNA binding protein variant D.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=83355;
RN [1]
RP SEQUENCE FROM N.A.
RA Havin L., Git A., Elisha Z., Oberman F., Yaniv K.,
RA Pressman Schwartz S., Standart N.M., Yisraeli J.K.;
RL Genes Dev. 0:0-0(1998).
RW [2]
SQ SEQUENCE FROM N.A.
RX MEDLINE=98228351; PubMed=9560341;
RA Deshler J.O., Highett M.I., Abranson T., Schnapp B.J.;
RT "A highly conserved RNA-binding protein for cytoplasmic mRNA
RT localization in vertebrates.";
RL Curr. Biol. 8:489-496(1998).
CC -|- SIMILARITY: CONTAINS 4 KH DOMAINS.
DR EMBL; AF054634; AAC18598.1; -
DR EMBL; AF055923; AAC41285.1; -
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR InterPro; IPR004087; KH dom.
DR InterPro; IPR004088; KH type 1.
DR Pfam; PF00013; KH; 4.
DR SMART; SM00322; KH; 4.
DR SMART; SM00360; RRM; 2.
DR PROSITE; PS00084; KH TYPE 1; 4.
DR PROSITE; PS0102; RRM; 2.
DR PROSITE; PS00030; RRM_RNP_1; FALSE NEG.
SQ SEQUENCE 594 AA; 65643 MW; 54CEA7BFF0856DD6 CRC64;

Alignment Scores:
Pred. No.: 3.75e-14 Length: 594
Score: 23.00 Matches: 23
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.05% Indels: 0
DB: 13 Gaps: 0

```

```

US-09-270-437D-6 (1-3412) x O73932 (1-594)

QY 1630 GCTGCGCGGTGATTGGCAAGAGTGGCAAGACCGTGAACGAACTCGAAGCTTAAACCACT 1689
Db |||||
QY 512 AlaGlyArgValIleGlyLysThrValAsnGluLeuGlnAsnLeuThrSer 531
Db |||||
QY 1690 GCAGAAAGTC 1698
Db |||||
QY 532 AlaGluVal 534
Db |||||

RESULT 6
O42254 PRELIMINARY; PRT; 576 AA.
AC O42254;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Zipcode-binding protein.
OS ZBPL.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97220007; PubMed=9121465;
RA Ross A.F., Oleynikov Y.S., Kisilevsky E.H., Taneja K.L., Singer R.H.;
RT "Characterization of a beta-actin mRNA zipcode-binding protein.";
RL Mol. Cell. Biol. 17:2158-2165(1997).
CC -|- SIMILARITY: CONTAINS 4 KH DOMAINS.
DR EMBL; AF026527; AAB82295.1; -
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR InterPro; IPR004087; KH dom.
DR InterPro; IPR004088; KH type 1.
DR InterPro; IPR00504; RNA_rec_mot.
DR Pfam; PF00013; KH; 4.
DR Pfam; PF00076; rrm; 2.
DR SMART; SM00322; KH; 4.
DR SMART; SM00360; RRM; 2.
DR PROSITE; PS00084; KH TYPE 1; 4.
DR PROSITE; PS0102; RRM; 2.
DR PROSITE; PS00030; RRM_RNP_1; FALSE NEG.
SQ SEQUENCE 576 AA; 63271 MW; 01AaF2D1D81C8811 CRC64;

Alignment Scores:
Pred. No.: 5.97e-10 Length: 576
Score: 19.00 Matches: 19
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.70% Indels: 0
DB: 13 Gaps: 0

US-09-270-437D-6 (1-3412) x O42254 (1-576)

QY 1630 GCTGCGCGGTGATTGGCAAGAGTGGCAAGACCGTGAACGAACTCGAAGCTTAAAC 1686
Db |||||
QY 498 AlaGlyArgValIleGlyLysThrValAsnGluLeuGlnAsnLeuThr 516
Db |||||

RESULT 7
Q9NZI8 PRELIMINARY; PRT; 577 AA.
AC Q9NZI8;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE mRNA-binding protein CRDBP.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.

```



RA Ioannidis P., Trangas T., Dimitriadis E., Samiotaki M.,  
RA Panoutsakopoulos G., Kyriazoglou I., Voutzoulis S., Tsiapalis C.M.,  
RA Kittas C., Agnantis N., Pandis N.;  
RT "Ectopic expression of a KH-domain containing protein, highly  
RT homologous to both human IMP-1 and mouse CRD-BP, in benign and  
RT malignant mesenchymal tumors.";  
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.  
CC -!- SIMILARITY: CONTAINS 4 KH DOMAINS.  
DR EMBL; AF198254; AAF37203.1; -.  
DR HSSP; P11940; 1CVJ.  
DR GO; GO:0003676; F:nucleic acid binding; IEA.  
DR InterPro; IPR004087; KH dom.  
DR InterPro; IPR004088; KH\_type\_1.  
DR InterPro; IPR000504; RNA\_rec\_mot.  
DR Pfam; PF00013; KH; 4.  
DR Pfam; PF00076; rrm; 2.  
DR SMART; SM00322; KH; 4.  
DR SMART; SM00360; RRM; 2.  
DR PROSITE; PS50084; KH TYPE 1; 4.  
DR PROSITE; PS50102; RRM; 2.  
SQ SEQUENCE 577 AA; 63456 MW; 0749A060F252D81D CRC64;

Alignment Scores:  
Pred. No.: 5.97e-10 Length: 577  
Score: 19.00 Matches: 19  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.70% Indels: 0  
DB: 4 Gaps: 0

US-09-270-437D-6 (1-3412) x Q3NZI8 (1-577)

QY 1489 AGATGGTCATCATCCGCGCACCGAAGCCAGTTCAGCCGACGGACGGATC 1545  
|||||  
Db 452 ArgMetValIleIleThrGlyProProGluAlaGlnPhelysAlaGlnGlyArgIle 470  
|||||

RESULT 8  
Q8CGX0  
ID Q8CGX0 PRELIMINARY; PRT; 577 AA.  
AC Q8CGX0;  
DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE B-actin zipcode binding protein 1.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Sprague-Dawley;  
RA Eom T., Singer R.H., Bassell G.J.;  
RT "Molecular interactions between rZBP1 and b-actin zipcode required for  
RT transport of mRNA and stimulation of spine growth.";  
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF541940; AA016210.1; -.  
DR GO; GO:0003676; F:nucleic acid binding; IEA.  
DR InterPro; IPR004087; KH dom.  
DR InterPro; IPR004088; KH type 1.  
DR InterPro; IPR000504; RNA\_rec\_mot.  
DR Pfam; PF00013; KH; 4.  
DR Pfam; PF00076; rrm; 2.  
DR SMART; SM00322; KH; 4.  
DR SMART; SM00360; RRM; 2.  
DR PROSITE; PS50084; KH TYPE 1; 4.  
DR PROSITE; PS50102; RRM; 2.  
SQ SEQUENCE 577 AA; 63436 MW; 0647676128FBD1EE CRC64;

Alignment Scores:  
Pred. No.: 5.97e-10 Length: 577  
Score: 19.00 Matches: 19  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0

Query Match: 1.70% Indels: 0  
DB: 11 Gaps: 0

US-09-270-437D-6 (1-3412) x Q8CGX0 (1-577)

QY 1630 GCTGCGCGGTGATTGGCAAGGTGCGAAGCCGTGCAAGAACTGCAGAACTTAACC 1686  
|||||  
Db 499 AlaGlyArgValIleGlyLysThrValAsnGluLeuGlnAsnLeuThr 517  
|||||

RESULT 9  
Q8BRH1  
ID Q8BRH1 PRELIMINARY; PRT; 577 AA.  
AC Q8BRH1;  
DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Insulin-like growth factor 2.  
GN IGF2BP1.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J;  
RX MEDLINE=22354683; PubMed=12466851;  
RA The FANTOM Consortium,  
RA the RIKEN Genome Exploration Research Group Phase I & II Team;  
RT "Analysis of the mouse transcriptome based on functional annotation of  
RT 60,770 full-length cDNAs.";  
RL Nature 420:563-573(2002).  
DR EMBL; AK044850; BAC32119.1; -.  
DR MGD; MGI:1890357; IGF2bp1.  
DR GO; GO:0003676; F:nucleic acid binding; IEA.  
DR InterPro; IPR004087; KH dom.  
DR InterPro; IPR004088; KH type 1.  
DR InterPro; IPR000504; RNA\_rec\_mot.  
DR Pfam; PF00013; KH; 4.  
DR Pfam; PF00076; rrm; 2.  
DR SMART; SM00322; KH; 4.  
DR SMART; SM00360; RRM; 2.  
DR PROSITE; PS50084; KH TYPE 1; 4.  
DR PROSITE; PS50102; RRM; 2.  
SQ SEQUENCE 577 AA; 63378 MW; D75897450841532E CRC64;

Alignment Scores:  
Pred. No.: 5.97e-10 Length: 577  
Score: 19.00 Matches: 19  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.70% Indels: 0  
DB: 11 Gaps: 0

US-09-270-437D-6 (1-3412) x Q8BRH1 (1-577)

QY 1630 GCTGCGCGGTGATTGGCAAGGTGCGAAGCCGTGCAAGAACTGCAGAACTTAACC 1686  
|||||  
Db 499 AlaGlyArgValIleGlyLysThrValAsnGluLeuGlnAsnLeuThr 517  
|||||

RESULT 10  
Q80US9  
ID Q80US9 PRELIMINARY; PRT; 577 AA.  
AC Q80US9;  
DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Igf2bp1 protein.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.



Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.70% Indels: 0  
DB: 11 Gaps: 0

US-09-270-437D-6 (1-3412) x 088477 (1-577)

QY 1630 GCTGGCGGGTGATTGGCAAGTGGCAGACCGTGAACGCTGAGAACTGAGAACTTAACC 1686

Db 499 AlaglyArgValileGlyLysGlyGlyLysThrValAsnGluLeuGlnAsnLeuThr 517

RESULT 12

ID 000425 PRELIMINARY; PRT; 579 AA.

AC 000425;

DT 01-JUL-1997 (TrEMBLrel. 04, Created)

DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DE Putative RNA binding protein KOC (KOC).

GN KOC.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI\_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RA Mueller-Pillasch F., Lacher U., Wallrapp C., Et AL.;

RL Oncogene 0:0-0(0).

RN [2]

RP SEQUENCE FROM N.A.

RC TISSUE=Pancreas;

RA Mueller-Pillasch F., Lacher U., Wallrapp C.;

RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.

CC -1- SIMILARITY: CONTAINS 4 KH DOMAINS.

DR EMBL; U97188; AAC35208.1; --

DR EMBL; U76705; AAD09223.1; --

DR GO; GO:0005737; C:cytoplasm; TAS.

DR GO; GO:0003723; F:RNA binding; TAS.

DR GO; GO:0007345; P:embryogenesis and morphogenesis; TAS.

DR GO; GO:0006412; P:protein biosynthesis; TAS.

DR GO; GO:0006396; P:RNA processing; TAS.

DR InterPro; IPR004087; KH dom.

DR InterPro; IPR004088; KH\_type\_1.

DR InterPro; IPR000504; RNA\_rec\_mot.

DR Pfam; PF00013; KH; 4.

DR Pfam; PF00076; xrm; 2.

DR SMART; SM00322; KH; 4.

DR SMART; SM00360; RRM; 2.

DR PROSITE; PS50084; KH\_TYPE\_1; 4.

DR PROSITE; PS50102; RRM; 2.

DR PROSITE; PS00030; RRM\_RNP\_1; FALSE NEG.

SQ SEQUENCE 579 AA; 63720 MW; AE5C3A8EE3C135C5 CRC64;

Alignment Scores:

Pred. No.: 5 97e-10 Length: 579  
Score: 19.00 Matches: 19  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.70% Indels: 0  
DB: 4 Gaps: 0

US-09-270-437D-6 (1-3412) x 000425 (1-579)

QY 1489 AGGATGTCATCATCCGGCCACCGGACCGGACCGGACCGGACCGGATC 1545

Db 452 ArgMetValilelleThrGlyProProGluAlaGlnPheLysAlaGlnGlyArgIle 470

RESULT 13

Q8C2J9

ID Q8C2J9 PRELIMINARY; PRT; 579 AA.

AC Q8C2J9;

DT 01-MAR-2003 (TrEMBLrel. 23, Created)

DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Insulin-like growth factor 2.  
GN IGF2BP3.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=NOD; TISSUE=Thymus;  
RX MEDLINE=22354683; PubMed=12466851;  
RA The FANTOM Consortium.  
RA the RIKEN Genome Exploration Research Group Phase I & II Team;  
RT "Analysis of the mouse transcriptome based on functional annotation of  
RL Nature 420:563-573(2002)."  
DR EMBL; AK088465; BAC40370.1; --  
DR MGD; MGI:1890359; Igf2bp3.  
DR GO; GO:0003676; F:nucleic acid binding; IEA.  
DR InterPro; IPR004087; KH\_dom.  
DR InterPro; IPR004088; KH\_type\_1.  
DR InterPro; IPR000504; RNA\_rec\_mot.  
DR Pfam; PF00013; KH; 4.  
DR Pfam; PF00076; xrm; 2.  
DR SMART; SM00322; KH; 4.  
DR SMART; SM00360; RRM; 2.  
DR PROSITE; PS50084; KH\_TYPE\_1; 4.  
DR PROSITE; PS50102; RRM; 2.  
SQ SEQUENCE 579 AA; 63551 MW; 937E601A95D06B77 CRC64;

Alignment Scores:

Pred. No.: 5 97e-10 Length: 579  
Score: 19.00 Matches: 19  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.70% Indels: 0  
DB: 11 Gaps: 0

US-09-270-437D-6 (1-3412) x Q8C2J9 (1-579)

QY 1489 AGGATGTCATCATCCGGCCACCGGACCGGACCGGACCGGACCGGATC 1545

Db 452 ArgMetValilelleThrGlyProProGluAlaGlnPheLysAlaGlnGlyArgIle 470

RESULT 14

Q9CFN8

ID Q9CFN8 PRELIMINARY; PRT; 579 AA.

AC Q9CFN8;

DT 01-JUN-2001 (TrEMBLrel. 17, Created)

DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DE 10 days embryo cDNA, RIKEN full-length enriched library,

DE clone:2610036B18, full insert sequence (Igf2 mRNA-binding protein 3)

DE (Insulin-like growth factor 2, binding protein 3).

GN IGF2BP3 OR 2610101N1RIK OR MIMP3.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

OX NCBI\_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Embryo;

RX MEDLINE=21085660; PubMed=11217851;

RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,

RA Akawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,

RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,

RA Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,

RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,

RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,

RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,

RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,

RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,

RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,

RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
RA Lyons P., Marchionni L., Mashima J., Marzarelli J., Mombaerts P.,  
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
RA Suzuki H., Toyoko-Oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,  
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,  
RA Hayashizaki Y.,  
RT "Functional annotation of a full-length mouse cDNA collection.";  
RL Nature 409:685-690(2001).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Mori H., Sakakibara S., Imai T., Nakamura Y., Iijima T., Suzuki A.,  
RA Yuasa Y., Takeda M., Okano H.,  
RT "Expression of mouse igf2 mRNA-binding protein 3 and its implications  
RT for the developing central nervous system.";  
RL J. Neurosci. Res. 0:0-0(2001).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6; TISSUE=Brain, and Olfactory epithelium;  
RX MEDLINE=22388257; PubMed=12477932;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
RA Krzywinski M.I., Skalska U., Smalls D.E., Schnerch A., Schein J.E.,  
RA Jones S.J., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [4]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6; TISSUE=Brain, and Olfactory epithelium;  
RA Strausberg R.;  
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.  
CC -!- SIMILARITY: CONTAINS 4 KH DOMAINS.  
DR EMBL; AK011689; BAB27779.1; -;  
DR EMBL; AB046173; BAB19755.1; -;  
DR EMBL; BC045138; AAH45138.1; -;  
DR EMBL; BC049082; AAH49082.1; -;  
DR MGD; MGI:1890359; Igf2bp3.  
DR GO; GO:0003676; F:nucleic acid binding; IEA.  
DR InterPro; IPR004087; KH\_dom.  
DR InterPro; IPR004088; KH\_type\_1.  
DR InterPro; IPR000504; RNA\_rec\_mot.  
DR Pfam; PF00013; KH; 4.  
DR Pfam; PF00076; rim; 2.  
DR SMART; SM00322; KH; 4.  
DR SMART; SM00360; RRM; 2.  
DR PROSITE; PSS0084; KH\_TYPE\_1; 4.  
DR PROSITE; PS0102; RRM; 2.  
SQ SEQUENCE 579 AA; 63574 MW; CABD9A435B392B7 CRC64;

Alignment Scores:  
Pred. No.: 5, 97e-10 Length: 579  
Score: 19.00 Matches: 19  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.70% Indels: 0  
DB: 11 Gaps: 0

US-09-270-437D-6 (1-3412) x Q9CPN8 (1-579)

OY 1489 AGGATGGTCATCATCACCGGCCACCGAGCCAGTTCAGGCCCGGCGATC 1545  
|||  
DB 452 ArgMetValIleIleThrGlyProGluAlaGlnPhelysAlaGlnGlyArgile 470  
|||  
RESULT 15  
Q7TP50  
ID Q7TP50 PRELIMINARY; PRT; 169 AA.  
AC Q7TP50;  
DT 01-OCT-2003 (TRENBLrel. 25, Created)  
DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)  
DE Ab2-255.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Xu C.S., Li W.Q., Li Y.C., Yang K.J., Yan H.M., Chang C.F., Zhao L.F.,  
RA Ma H., Wang L., Wang S.F., Han H.P., Wang G.P., Chai L.Q., Yuan J.Y.,  
RA Shi J.B., Rahman S., Wang Q.N., Zhang J.B.;  
RT "Liver regeneration after PH";  
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY325199; AAP92600.1; -;  
SQ SEQUENCE 169 AA; 17743 MW; 19BF6295C1000CBA CRC64;

Alignment Scores:  
Pred. No.: 0.000123 Length: 169  
Score: 14.00 Matches: 14  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.25% Indels: 0  
DB: 11 Gaps: 0

US-09-270-437D-6 (1-3412) x Q7TP50 (1-169)

OY 1489 AGGATGGTCATCATCACCGGCCACCGAGCCAGTTCAG 1530  
|||  
DB 61 ArgMetValIleIleThrGlyProGluAlaGlnPhelys 74  
|||

Search completed: July 16, 2004, 11:29:33  
Job time : 183.5 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 21, 2004, 12:27:22 ; Search time 8233 Seconds  
(without alignments)  
12375.776 Million cell updates/sec

Title: US-09-270-437D-6  
Perfect score: 3412  
Sequence: 1 ggcagcggaggaggagga.....aaccttgaaaatgtttattt 3412

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 27513289 seqs, 14931090276 residues

Word size : 0

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : EST.\*  
1: em\_estba.\*  
2: em\_esthum.\*  
3: em\_estin.\*  
4: em\_estnu.\*  
5: em\_estov.\*  
6: em\_estpl.\*  
7: em\_estro.\*  
8: em\_htc.\*  
9: gb\_est1.\*  
10: gb\_est2.\*  
11: gb\_htc.\*  
12: gb\_est3.\*  
13: gb\_est4.\*  
14: gb\_est5.\*  
15: em\_estfun.\*  
16: em\_estom.\*  
17: em\_gss\_hum.\*  
18: em\_gss\_inv.\*  
19: em\_gss\_pln.\*  
20: em\_gss\_vrt.\*  
21: em\_gss\_fun.\*  
22: em\_gss\_man.\*  
23: em\_gss\_mus.\*  
24: em\_gss\_pro.\*  
25: em\_gss\_rod.\*  
26: em\_gss\_phg.\*  
27: em\_gss\_vrl.\*  
28: gb\_gss1.\*  
29: gb\_gss2.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	834	24.4	874	13 BQ691878	BQ691878 AGENCOURT
2	815	23.9	905	12 BM453327	BM453327 AGENCOURT
3	771	22.6	832	13 BX644668	BX644668 DKFZp781G
4	747	21.9	851	13 BQ221568	BQ221568 AGENCOURT

5	734	21.5	940	14	CD243351	AGENCOURT
6	715	21.0	824	14	CF593505	AGENCOURT
7	702	20.6	1089	12	BM806132	AGENCOURT
8	698	20.5	999	12	BM561242	AGENCOURT
9	670	19.6	724	12	BM977219	UI-CF-DU1
10	642	18.8	836	13	BQ212500	AGENCOURT
11	634	18.6	647	13	BU674592	UI-CF-DU0
12	617	18.1	889	12	BG748346	AGENCOURT
13	594	17.4	781	12	BG575889	AGENCOURT
14	590	17.3	894	13	BU156245	AGENCOURT
15	585	17.1	918	13	BQ212353	AGENCOURT
16	584	17.1	584	13	BX642777	DKFZp781I
17	566	16.6	762	13	BU601634	AGENCOURT
18	525	15.4	628	12	BI862502	AGENCOURT
19	507	14.9	939	13	BQ961314	AGENCOURT
20	505	14.8	607	12	BM724394	UI-E-E01
21	489	14.3	612	13	BG679082	UI-CF-DU1
22	480	14.1	531	10	BG009869	QV1-GN032
23	477	14.0	620	10	BE914566	IL3-UT011
24	448	13.1	729	10	BE278075	601179604
25	440	12.9	918	14	CF995222	AGENCOURT
26	440	12.9	981	13	BQ887769	AGENCOURT
27	438	12.8	453	9	AL121466	DKFZp762L
28	434	12.7	594	12	BI858380	603383965
29	423	12.4	459	10	BE277048	601178938
30	416	12.2	638	12	BM843467	K-EST0121
31	409	12.0	449	13	BQ361219	PMO-OT023
32	397	11.6	485	13	BQ355956	PMO-HT119
33	397	11.6	485	13	BQ355967	PMO-HT116
34	380	11.1	759	10	BE622021	601440673
35	378	11.1	720	13	BX101497	601440673
36	376	11.0	976	10	BQ025065	602276408
37	374	11.0	386	9	AV698363	AV698363
38	365	10.7	845	13	BU155790	AGENCOURT
39	357	10.5	469	13	BX282561	602282561
40	346	10.1	1212	13	BU902212	AGENCOURT
41	343	10.1	344	9	AI004702	ot-95f07.x
42	341	10.0	506	10	BE928277	QV4-CT049
43	326	9.6	1078	12	BM462051	AGENCOURT
44	325	9.5	526	10	BE896980	601439521
45	310	9.1	608	12	BG254615	602368564

ALIGNMENTS

RESULT 1  
BQ691878  
LOCUS  
DEFINITION BQ691878 874 bp mRNA linear EST 15-JUL-2002  
AGENCOURT\_8034698 NIH\_MGC\_110 Homo sapiens cDNA clone IMAGE:6207026  
5', mRNA sequence.  
ACCESSION BQ691878  
VERSION BQ691878.1 GI:21817194  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 874)  
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-rc@mail.nih.gov](mailto:cgapbs-rc@mail.nih.gov)  
Tissue Procurement: ATCC  
cDNA Library Preparation: Rubin Laboratory  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LLCM2362 row: h column: 03  
High quality sequence stop: 653.

FEATURES  
sourceLocation/Qualifiers  
1..874

/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6207026"  
/tissue\_type="ductal carcinoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH MGC 110"  
/note="Organ: pancreas; Vector: pOTB7; Site: 1: XhoI;  
Site 2: EcoRI; cDNA made by oligo-dT priming.  
Directionally cloned into EcoRI/XhoI sites using the  
following 5' adaptor: GGACAG(G). Library constructed by  
Ling Hong in the laboratory of Gerald M. Rubin (University  
of California, Berkeley) using ZAP-cDNA synthesis kit  
(Stratagene) and Superscript II RT (Life Technologies).  
Note: this is a NIH\_MGC Library."

## ORIGIN

Query Match 24.4%; Score 834; DB 13; Length 874;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 834; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1824 GGAGCAGAAATACCTCTAGGAGTGCCTCAGCGCAGCAAGTGAAGCTCCACAGCA 1883  
DB 1 GGAGCAGAAATACCTCTAGGAGTGCCTCAGCGCAGCAAGTGAAGCTCCACAGCA 60

QY 1884 CCAGCAAAACAGGATGAATGTAGCCCTTCCACACCTGACAGATGAGACCAACGCA 1943  
DB 61 CCAGCAAAACAGGATGAATGTAGCCCTTCCACACCTGACAGATGAGACCAACGCA 120

QY 1944 GCCAGCCAGATCGGAGCAAAACCAAGACCATCTGAGGAATGAGAAGTCTCGGAGGCGG 2003  
DB 121 GCCAGCCAGATCGGAGCAAAACCAAGACCATCTGAGGAATGAGAAGTCTCGGAGGCGG 180

QY 2004 CCAGGACCTCTGCCAGGCCCTGAGAACCCAGGGCCGAGAGGGGGGGAAGGTGAG 2063  
DB 181 CCAGGACCTCTGCCAGGCCCTGAGAACCCAGGGCCGAGAGGGGGGGAAGGTGAG 240

QY 2064 CCAGGTTTGCAGAACCAACCGAGCCCGCCCTCCGCGCCCGCCAGGGCTTCTCAGGCTTCA 2123  
DB 241 CCAGGTTTGCAGAACCAACCGAGCCCGCCCTCCGCGCCCGCCAGGGCTTCTCAGGCTTCA 300

QY 2124 GCCATCCACTTCAACATCCACTCGGATCTCTCTGAATCCCAAGCAGCTATCCCTTTTA 2183  
DB 301 GCCATCCACTTCAACATCCACTCGGATCTCTCTGAATCCCAAGCAGCTATCCCTTTTA 360

QY 2184 GTTGAACCTAACATAGTGAAGTGTCTCAAGCCAGCAAGCAATGCAACCCCTTTTCTGTG 2243  
DB 361 GTTGAACCTAACATAGTGAAGTGTCTCAAGCCAGCAAGCAATGCAACCCCTTTTCTGTG 420

QY 2244 GCAAAATCGTCTGTACATGTGTGTACATATTAGAAAGGGAAGATGTTAAGATATGTGGC 2303  
DB 421 GCAAAATCGTCTGTACATGTGTGTACATATTAGAAAGGGAAGATGTTAAGATATGTGGC 480

QY 2304 CTGTGGGTTACACAGGGTGCCTGACGGGTAAATATTTTGAATAATATATATCAATATAA 2363  
DB 481 CTGTGGGTTACACAGGGTGCCTGACGGGTAAATATTTTGAATAATATATATCAATATAA 540

QY 2364 CTCACCTAACCTCAATTTTAAATCAATTTAAATTTTCTTTTCTTTTAAAGAGAGCA 2423  
DB 541 CTCACCTAACCTCAATTTTAAATCAATTTAAATTTTCTTTTCTTTTAAAGAGAGCA 600

QY 2424 GGCTTTTCTAGACTTTTAAAGATAAAGTCTTTTGGGAGGTCTCAGCGGTGTAGAGAGGCT 2483  
DB 601 GGCTTTTCTAGACTTTTAAAGATAAAGTCTTTTGGGAGGTCTCAGCGGTGTAGAGAGGCT 660

QY 2484 TTGAGGCCACCCGACAAAATTCACCCAGAGGGAATCTCGTCCGAGGACACTCAACGGC 2543  
DB 661 TTGAGGCCACCCGACAAAATTCACCCAGAGGGAATCTCGTCCGAGGACACTCAACGGC 720

QY 2544 AGTTCTGATCACTGTGTATGTCAACAGAGGATACCGTCTCTCTTGAAGAGAACTC 2603

Db 721 AGTTCTGATCACCTGTGTATGTCAACAGAGGATACCGTCTCTCTTGAAGAGAACTC 780

QY 2604 TGTCACTCTCATGCTGTCTAGCTCATACCCCATTTCTCTTCTTCTTCAAGG 2657  
Db 781 TGTCACTCTCATGCTGTCTAGCTCATACCCCATTTCTCTTCTTCTTCAAGG 834

RESULT 2  
EM453327  
LOCUS BM453327 905 bp mRNA linear EST 05-FEB-2002  
DEFINITION AGENCOURT 6397801 NIH\_MGC\_72 Homo sapiens cDNA clone IMAGE:5527679  
5', mRNA sequence.  
ACCESSION BM453327  
VERSION BM453327.1 GI:18502367  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 905)  
AUTHORS NIH-MGC http://mgc.nci.nih.gov/  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-re@mail.nih.gov  
Tissue Procurement: ATCC/DCTD/DTP  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLAM12203 Row: e column: 24  
High quality sequence stop: 568.

FEATURES  
source  
Location/Qualifiers  
1..905  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5527679"  
/tissue\_type="melanotic melanoma"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH MGC 72"  
/note="Organ: skin; Vector: pCMV-SPORT6; Site 1: NotI;  
Site 2: SalI; Cloned unidirectionally. Primer: oligo dT.  
Average insert size 2 kb. Library constructed by Life  
Technologies."

## ORIGIN

Query Match 23.9%; Score 815; DB 12; Length 905;  
Best Local Similarity 100.0%; Pred. No. 6.8e-312; Indels 0; Gaps 0;  
Matches 815; Conservative 0; Mismatches 0;

QY 492 CGGCGATCAGTTTGAGAACTACTCCTTCAAGATTCTACATCCCGATGAAGAGTGAG 551  
DB 1 CGGCGATCAGTTTGAGAACTACTCCTTCAAGATTCTACATCCCGATGAAGAGTGAG 60

QY 552 CTCCTCTCCGCTCAGCGAGCCGCGTGGGACACTCTTCCCGGGAGCAAGGCCA 611  
DB 61 CTCCTCTCCGCTCAGCGAGCCGCGTGGGACACTCTTCCCGGGAGCAAGGCCA 120

QY 612 CGCCCTCTGGGGGCACTTCTCAGGCCAGACAGATTGATTTCCCGCTGGGATCCTGTGCC 671  
DB 121 CGCCCTCTGGGGGCACTTCTCAGGCCAGACAGATTGATTTCCCGCTGGGATCCTGTGCC 180

QY 672 CACCGATTTGTTGGTCCCATCATCGGAAAGAGGGCTTGACCATATAAGAACATCACTAA 731  
DB 181 CACCGATTTGTTGGTCCCATCATCGGAAAGAGGGCTTGACCATATAAGAACATCACTAA 240

QY 732 GCAGACCCAGTCCCGGTAGTATCCATAGAAAGAGAACTCTCGAGCTGCAGAGAGCC 791  
DB 241 GCAGACCCAGTCCCGGTAGTATCCATAGAAAGAGAACTCTCGAGCTGCAGAGAGCC 300

```
QY 792 TGTCACCATCCATGCCACCCAGAGGGGACTTCTGAAGCATGCCGATGATTTCTTGAAT 851
Db 301 TGTCACCATCCATGCCACCCAGAGGGGACTTCTGAAGCATGCCGATGATTTCTTGAAT 360
QY 852 CATGCAGAAAGAGGCAGATGAGACCAAACTAGCCGAAGAGATTCCTCTGAAATCTTTGGC 911
Db 361 CATGCAGAAAGAGGCAGATGAGACCAAACTAGCCGAAGAGATTCCTCTGAAATCTTTGGC 420
QY 912 ACACAATGCTTGGTTGGAGACTGATTTGGAAAAGAGGAGGAGAAATTTGAGAAAATTGA 971
Db 421 ACACAATGCTTGGTTGGAGACTGATTTGGAAAAGAGGAGGAGAAATTTGAGAAAATTGA 480
QY 972 ACATGAAACAGGACCAAGATACAACTCTATCTTTGAGGATTTGAGCATATACAAACC 1031
Db 481 ACATGAAACAGGACCAAGATACAACTCTATCTTTGAGGATTTGAGCATATACAAACC 540
QY 1032 GGAAGAACCATCACTGTGAAGGGCACAGTTGAGGCTGTGCCAGTGTGAGATAGAT 1091
Db 541 GGAAGAACCATCACTGTGAAGGGCACAGTTGAGGCTGTGCCAGTGTGAGATAGAT 600
QY 1092 TATGAAGAGCTCGTGAGGCTTTGAAAATGATATGCTGGCTGTAAACCAAGCCAA 1151
Db 601 TATGAAGAGCTCGTGAGGCTTTGAAAATGATATGCTGGCTGTAAACCAAGCCAA 660
QY 1152 TCTGATCCAGAGGTTGAACCTCAGCGCACCTTGGCATCTTTTCAACAGGACTGTCCGTGCT 1211
Db 661 TCTGATCCAGAGGTTGAACCTCAGCGCACCTTGGCATCTTTTCAACAGGACTGTCCGTGCT 720
QY 1212 ATCTCCACGACGAGGCCCCGGAGCTCCCCCGCTGTCGCCCTTACCACCCCTTCACTAC 1271
Db 721 ATCTCCACGACGAGGCCCCGGAGCTCCCCCGCTGTCGCCCTTACCACCCCTTCACTAC 780
QY 1272 CCACTCGGATCTTCTCCAGCCTGTACCCCATC 1306
Db 781 CCACTCGGATCTTCTCCAGCCTGTACCCCATC 815
```

```
RESULT 3
BX644668
LOCUS
DEFINITION
  BX644668 BX644668.1 GI:34479001
  DKFZp781G0234 r1 781 (synonym: hlcc4) Homo sapiens cDNA clone
  DKFZp781G0234 5', mRNA sequence.
ACCESSION
  BX644668
VERSION
  BX644668.1
KEYWORDS
  EST.
SOURCE
  Homo sapiens (human)
  Homo sapiens
  Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
  1 (bases 1 to 832)
  Lauber,J., Bahr,A., Mewes,H.W., Weil,B., Amid,C., Osanger,A.,
  Fobo,G., Han,M. and Wiemann,S.
  EST (Lauber,J., Bahr,A., Mewes,H.W., Weil,B., Amid,C., et al.)
  Unpublished (2003)
  Contact: MIPS
  MIPS
  Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
  This is the 5' sequence of the clone insert
  Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
  Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
  sequenced by Qiagen (Hilden/Germany) within the cDNA sequencing
  consortium of the German Genome Project.
  No sl sequence available.
  This clone (DKFZp781G0234) is available at the RZPD in Berlin.
  Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
  Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
```

```
FEATURES
  source
    1..832
    Location/Qualifiers
      /organism="Homo sapiens"
      /mol_type="mRNA"
      /db_xref="taxon:9606"
      /clone="DKFZp781G0234"
      /dev_stage="adult"
```

```
ORIGIN
Query Match 22.6%; Score 771; DB 13; Length 832;
Best Local Similarity 100.0%; Pred. No. 1.9e-294; Indels 0; Gaps 0;
Matches 771; Conservative 0; Mismatches 0;
/lab_host="DH10B"
/clone_lib="781 (synonym: hlcc4)"
/note="Vector: pSport1_Sfi; Site_1: SfiIA; Site_2: SfiIB;
cDNA-collection"
QY 1732 GAGGAGTATCGTCTCAGAAATTTATCGGGCAGCTTTCTTTGCTAGCCAGACTGACAGCGCAAG 1791
Db 13 GAGGAGTATCGTCTCAGAAATTTATCGGGCAGCTTTCTTTGCTAGCCAGACTGACAGCGCAAG 72
QY 1792 ATCAGGGAATTTGTACAAACAGGTGAAGCAGGAGCAGAAATACCTCAGGAGTCGCC 1851
Db 73 ATCAGGGAATTTGTACAAACAGGTGAAGCAGGAGCAGAAATACCTCAGGAGTCGCC 132
QY 1852 TCACAGCGCAGCAAGTGAAGGCTCCACAGGCAACCAACAAACAAACGATGAATGTAGCCC 1911
Db 133 TCACAGCGCAGCAAGTGAAGGCTCCACAGGCAACCAACAAACGATGAATGTAGCCC 192
QY 1912 TTCCAAACACTGACAGAATGAGACCAACGAGCCAGAGTCGGGAGCAAAACCAAGA 1971
Db 193 TTCCAAACACTGACAGAATGAGACCAACGAGCCAGAGTCGGGAGCAAAACCAAGA 252
QY 1972 CCATCTGAGGAATGAGAAGTCTGCGAGGCGGCGAGGACTCTGCGAGGCGCTTGAGAAC 2031
Db 253 CCATCTGAGGAATGAGAAGTCTGCGAGGCGGCGAGGACTCTGCGAGGCGCTTGAGAAC 312
QY 2032 CCAGAGGCGCGAGGCGGCGGAGGTCAGCAGGTTTGCAGAACCAACCGAGCCCCG 2091
Db 313 CCAGAGGCGCGAGGCGGCGGAGGTCAGCAGGTTTGCAGAACCAACCGAGCCCCG 372
QY 2092 CTTCCCGCCCCCGAGGCTTCTGAGGCTTACGCCATCCACTCCACTCGGATC 2151
Db 373 CTTCCCGCCCCCGAGGCTTCTGAGGCTTACGCCATCCACTCCACTCGGATC 432
QY 2152 TCTCTCTGAATCCACGACGCTATCCCTTTAGTTGAACATAGGTGAACGTGTCA 2211
Db 433 TCTCTCTGAATCCACGACGCTATCCCTTTAGTTGAACATAGGTGAACGTGTCA 492
QY 2212 AGCCCAAGCAAAATGCACACCCCTTTTCTGTGGCAATCGTCTCTGATGTGTGACA 2271
Db 493 AGCCCAAGCAAAATGCACACCCCTTTTCTGTGGCAATCGTCTCTGATGTGTGACA 552
QY 2272 TATTAGAAAGGGAAGATGTTAAGATATGTGGCTGTGGGTTACACAGGGTGCCTGACGG 2331
Db 553 TATTAGAAAGGGAAGATGTTAAGATATGTGGCTGTGGGTTACACAGGGTGCCTGACGG 612
QY 2332 GTAAATATTTTGAATAATATATCAATACTCAACTCACTCACTCACTCACTCACTCACT 2391
Db 613 GTAAATATTTTGAATAATATATCAATACTCAACTCACTCACTCACTCACTCACTCACT 672
QY 2392 ATTAATTTTTTTTTTTCTTTTAAAGAGAAAGCAGGCTTTTCTAGACTTTTAAAGATAAAGT 2451
Db 673 ATTAATTTTTTTTTTTCTTTTAAAGAGAAAGCAGGCTTTTCTAGACTTTTAAAGATAAAGT 732
QY 2452 CTTTGGGAGGCTCTCAGCGTGTAGAGGAGCTTTTGGAGGCCACCCGACACAA 2502
Db 733 CTTTGGGAGGCTCTCAGCGTGTAGAGGAGCTTTTGGAGGCCACCCGACACAA 783
```

```
RESULT 4
BX221568
LOCUS
DEFINITION
  BX221568 BX221568.1 GI:20402968
  AGENCOURT_7559207 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:6046477
  5', mRNA sequence.
ACCESSION
  BX221568
VERSION
  BX221568.1
KEYWORDS
  EST.
SOURCE
  Homo sapiens (human)
  Homo sapiens
  ORGANISM
```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheraia; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 851)  
NIH-MGC <http://mgs.nci.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: ATCC/DCTD/DTF  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LLAM13292 row: f column: 14  
High quality sequence stop: 637.  
Location/Qualifiers  
1..851

FEATURES  
source

/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6046477"  
/tissue\_type="melanotic melanoma"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH MGC 72"  
/note="Organ: skin; Vector: pCMV-SPORT6; Site 1: NotI;  
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.  
Average insert size 2 kb. Library constructed by Life  
Technologies."

ORIGIN

Query Match 21.9%; Score 747; DB 13; Length 851;  
Best Local Similarity 99.9%; Pred. No. 5.9e-285;  
Matches 797; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 47 GCGGGCTTCGCGGAAGAGACGGATGATGAACAAGCTTTACATCGGAACCTGAGCCCGC 106  
Db 1 GCGGGCTTCGCGGAAGAGACGGATGATGAACAAGCTTTACATCGGAACCTGAGCCCGC 60  
QY 107 CCGTCACCGCGACGACCTCGGCGAGCTCTTTGGGGACAGGAAGTGCCTCGCGGGAC 166  
Db 61 CCGTCACCGCGACGACCTCGGCGAGCTCTTTGGGGACAGGAAGTGCCTCGCGGGAC 120  
QY 167 AGGTCCTGCTCAAGTCCGGCTACGCTTCGTTGGACTCCCGACACAGAACTGGCCATCC 226  
Db 121 AGGTCCTGCTCAAGTCCGGCTACGCTTCGTTGGACTCCCGACACAGAACTGGCCATCC 180  
QY 227 GCGGCATCGAGACCTCTCGGGTAAAGTGAATTCATGGGAAATCATGGAAGTTGATT 286  
Db 181 GCGGCATCGAGACCTCTCGGGTAAAGTGAATTCATGGGAAATCATGGAAGTTGATT 240  
QY 287 ACTCAGTCTTAAAGCTAAGGACGAGGAGAAATTCAGATTCGAAACATCCCTCTCACC 346  
Db 241 ACTCAGTCTTAAAGCTAAGGACGAGGAGAAATTCAGATTCGAAACATCCCTCTCACC 300  
QY 347 TGCACTGGAGGTGTGATGCACTTTTGGCTCAATATGGGACAGTGGAGATGTGAAC 406  
Db 301 TGCACTGGAGGTGTGATGCACTTTTGGCTCAATATGGGACAGTGGAGATGTGAAC 360  
QY 407 AAGTCAACACAGACAGAAACCGCGTTGTCAACGTGCATATGCAACAGAGAAG 466  
Db 361 AAGTCAACACAGACAGAAACCGCGTTGTCAACGTGCATATGCAACAGAGAAG 420  
QY 467 CAAAAATAGCCATGAGAGCTAAGCGGCGATCAGTTTGAACTACTCTTCAAGATTT 526  
Db 421 CAAAAATAGCCATGAGAGCTAAGCGGCGATCAGTTTGAACTACTCTTCAAGATTT 480  
QY 527 CCTACATCCCGGATGAAGAGGTGAGCTCCCTTCGCCCTTCAGCGAGCCAGCGTGGG 586  
Db 481 CCTACATCCCGGATGAAGAGGTGAGCTCCCTTCGCCCTTCAGCGAGCCAGCGTGGG 540  
QY 587 ACCTACTCTTCCCGGAGCAAGGCCACCGCCCTCGGGGGCACTTCTCAGGCCACAGAGATTG 646

Db 541 ACCACTCTTCCGGGAGCAAGCCACGCTCTGGGGGCACTTCTCAGGCCACAGAGATTG 600  
QY 647 ATTTCCTCGCTCGGATCCTGTCCTCCACCCAGTTTGTGGTCCCATCATCGGAAAGGAGG 706  
Db 601 ATTTCCTCGCTCGGATCCTGTCCTCCACCCAGTTTGTGGTCCCATCATCGGAAAGGAGG 660  
QY 707 GCTTGACCATTAAGAACATCACTAAGCAGACCCAGTCCCGGTTAGATATCCATAGAAAG 766  
Db 661 GCTTGACCATTAAGAACATCACTAAGCAGACCCAGTCCCGGTTAGATATCCATAGAAAG 720  
QY 767 AGAACTCTGGAGCTGCAGAGAAAGCTGTCCATCCATCCACCCAGAGGGGACTTCTG 826  
Db 721 AGAACTCTGGAGCTGCAGAGAAAGCTGTCCATCCATCCATCCACCCAGAGGGGACTTCTG 780  
QY 827 AAGCATGCCGATGATTC 844  
Db 781 AAGCATGCCGATGATTC 798

RESULT 5

CD243351

LOCUS

DEFINITION

IMAGE:30383283 5', mRNA sequence.

CD243351

ACCESSION

CD243351.1 GI:31003815

EST.

KEYWORDS

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheraia; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 940)

NIH-MGC <http://mgs.nci.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

TITLE

JOURNAL

COMMENT

Email: cgapbs-remail.nih.gov

Tissue Procurement: Dr. Michael Brownstein

cDNA Library Preparation: Invitrogen Corp

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: NDAM446 row: o column: 04

High quality sequence start: 18

High quality sequence stop: 695.

Location/Qualifiers

1..940

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:30383283"

/lab\_host="DH10B-Ton A (T1 and T5 phage resistances)"

/clone\_lib="NIH MGC 180"

/note="Organ: Testis; Vector: pCMV-SPORT6.1; Site 1: NotI;

Site 2: EcoRV (destroyed); Library is oligo-dT primed and

directionally cloned (EcoRV site is destroyed upon

cloning). Average insert size 1.68 kb. Library was

constructed by (Invitrogen). Note: this is a NIH\_MGC

Library."

ORIGIN

Query Match 21.5%; Score 734; DB 14; Length 940;

Best Local Similarity 100.0%; Pred. No. 7.7e-280;

Matches 734; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 125 TCGGCGAGCTCTTTGGGGACAGGAAGCTGCCCTCGCGGGACAGGTCTCTGTAAGTCCG 184

Db 28 TCGGCGAGCTCTTTGGGGACAGGAAGCTGCCCTCGCGGGACAGGTCTCTGTAAGTCCG 87

QY 185 GCTACGCTTCTGTTGACTACCCGACAGAACTGGGCGCATCCGCGCATCGAGACCTCT 244



```
Db      88  GCTACGCTTCGTGGACTACCCGACCAAGACTGGGCCATCCGGCCATCGAGACCTCT 147
QY      245  CGGGTAAAGTGGAAATTCATGCGAAATCATGGAAGTTGATTACTCAGTCTCTAAAAAGC 304
Db      148  CGGGTAAAGTGGAAATTCATGCGAAATCATGGAAGTTGATTACTCAGTCTCTAAAAAGC 207
QY      305  TANGGAGCAGGAAATTCAGATTCGAACATCCCTCTCCTCAGTGCAGTGGAGGTGTGG 364
Db      208  TAAGGAGCAGGAAATTCAGATTCGAACATCCCTCTCCTCAGTGCAGTGGAGGTGTGG 267
QY      365  ATGGACTTTTGGCTCAATATGGGACAGTGGAGATTCGGAACAGTCAACACAGACACAG 424
Db      268  ATGGACTTTTGGCTCAATATGGGACAGTGGAGATTCGGAACAGTCAACACAGACACAG 327
QY      425  AAACCGCGTTTCAAGCTCACATATCAACAAGAGAGAAAGCAAAATAGCCATGGAGA 484
Db      328  AAACCGCGTTTCAAGCTCACATATCAACAAGAGAGAAAGCAAAATAGCCATGGAGA 387
QY      485  AGCTAAGCGGGCATCAGTTTGAGAACTACTCCTTCAGATTTCTACATCCCGATGAAG 544
Db      388  AGCTAAGCGGGCATCAGTTTGAGAACTACTCCTTCAGATTTCTACATCCCGATGAAG 447
QY      545  AGGTGAGTCCCCCTTCCGCCCTCAGCGAGCCAGCGTGGGACCACTCTTCCCGGAGC 604
Db      448  AGGTGAGTCCCCCTTCCGCCCTCAGCGAGCCAGCGTGGGACCACTCTTCCCGGAGC 507
QY      605  AAGGCCACGCCCTTGGGGGCACTTCTCAGCCACAGACAGATTGATTTCCCGTCCGATCC 664
Db      508  AAGGCCACGCCCTTGGGGGCACTTCTCAGCCACAGACAGATTGATTTCCCGTCCGATCC 567
QY      665  TGCTTCCACCCAGTTTGTGTGCTGCTATCCGAAAGAGGGCTTCACCATAAAGAA 724
Db      568  TGCTTCCACCCAGTTTGTGTGCTGCTATCCGAAAGAGGGCTTCACCATAAAGAA 627
QY      725  TCATTAAGCAGACCCAGTCCCGGGTAGATATCCATAGAAAGAGAACTCTGGAGCTCAG 784
Db      628  TCATTAAGCAGACCCAGTCCCGGGTAGATATCCATAGAAAGAGAACTCTGGAGCTCAG 687
QY      785  AGAAGCTGTACCATCCATCCATGCCACCCAGAGGGGATTTCTGAAGCATGCCATGATC 844
Db      688  AGAAGCTGTACCATCCATCCATGCCACCCAGAGGGGATTTCTGAAGCATGCCATGATC 747
QY      845  TTGAATCATGCAG 858
Db      748  TTGAATCATGCAG 761
```

```
RESULT 6
CF593505
LOCUS      CF593505      824 bp      mRNA      linear      EST 26-SEP-2003
DEFINITION AGENCOURT_15622249 NIH_MGC_147 Homo sapiens cDNA clone
IMAGE:30531076 5', mRNA sequence.
ACCESSION  CF593505
VERSION     CF593505.1 GI:36347108
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1. (bases 1 to 824)
AUTHORS    NIH-MGC http://mgi.nci.nih.gov/.
TITLE      National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL    Unpublished (1999)
COMMENT    Contact: Daniela S. Gerhard, Ph.D.
            Office of Cancer Genomics
            National Cancer Institute / NIH
            Bldg. 31 Rm10A07 Bethesda, MD 20892
            Email: cgapbs-remail.nih.gov
            Tissue Procurement: Dr. Stefan Hansson
            cDNA Library Preparation: Michael J. Brownstein (NHGRI) with help
            and advice from Piero Carninci (RIKEN)
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
```

DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
plate: NDAM621 row: e column: 05  
High quality sequence stop: 682.

FEATURES  
Location/Qualifiers  
source  
1..824  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:30531076"  
/tissue\_type="Human Placenta"  
/lab\_host="DH10B Tona"  
/clone\_lib="NIH\_MGC\_147"  
/notes="Organ: placenta; Vector: pBluescriptR; Site 1: all-XhoI; Site 2: BamHI; Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTVN-3', size-selected for average insert size 2.3 kb and normalized to ROT 5. This is a primary library enriched for full-length clones and constructed using the Cap-trapper method (Carninci, in preparation). Library constructed by M. Brownstein (NHGRI/NHGRI, National Institutes of Health). Note: This is a NIH\_MGC library."

ORIGIN

Query Match 21.0%; Score 715; DB 14; Length 824;  
Best Local Similarity 100.0%; Pred. No. 2.8e-272;  
Matches 715; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GG CAGCGGAGGAGCGGAGCGCGGTACCGGGCCGGGGAGCGCGGGCTCTCGGGG 60  
Db 38 GG CAGCGGAGGAGCGGAGCGCGGTACCGGGCCGGGGAGCGCGGGCTCTCGGGG 97  
QY 61 AA GAGACGGATGATGAACAAAGCTTTTACATCGGAAACCTGAGCCCCCGCGTCACGCGAC 120  
Db 98 AAGAGACGGATGATGAACAAAGCTTTTACATCGGAAACCTGAGCCCCCGCGTCACGCGAC 157  
QY 121 GACCTCGGAGCTCTTTGGGAGCAGAGAGCTGCCCTTGGCGGAGCAGGTCCTGCTGAAG 180  
Db 158 GACCTCGGAGCTCTTTGGGAGCAGAGAGCTGCCCTTGGCGGAGCAGGTCCTGCTGAAG 217  
QY 181 TCCGGCTACGCTTCTGTGGACTACCCGACCAAGCTGGGCCATCCGCGCCATCGAGACC 240  
Db 218 TCCGGCTACGCTTCTGTGGACTACCCGACCAAGCTGGGCCATCCGCGCCATCGAGACC 277  
QY 241 CTCTCGGGTAAAGTGGAAATTCATGGGAAATCATGGAAGTTGATTACTCAGTCTCTAAA 300  
Db 278 CTCTCGGGTAAAGTGGAAATTCATGGGAAATCATGGAAGTTGATTACTCAGTCTCTAAA 337  
QY 301 AAGCTAAGGAGCAGGAAATTCAGATTCGAAACATCCCTCTCCTCAGCTGCAGTGGAGGTG 360  
Db 338 AAGCTAAGGAGCAGGAAATTCAGATTCGAAACATCCCTCTCCTCAGCTGCAGTGGAGGTG 397  
QY 361 TTGGATGGACTTTTGGCTCAATATGGGACAGTGGAGATGTGGAACAAGTCAACACAGAC 420  
Db 398 TTGGATGGACTTTTGGCTCAATATGGGACAGTGGAGATGTGGAACAAGTCAACACAGAC 457  
QY 421 ACAGAAACCGCGTTGTCAACGTCACATATGCAACAGAGAGAGAGAGAGAGAGAGAGAG 480  
Db 458 ACAGAAACCGCGTTGTCAACGTCACATATGCAACAGAGAGAGAGAGAGAGAGAGAGAG 517  
QY 481 GAGAAGCTAAGCGGGATCAGTTTGGAGACTACTCTTCAAGATTTCTTACATCCCGGAT 540  
Db 518 GAGAAGCTAAGCGGGATCAGTTTGGAGACTACTCTTCAAGATTTCTTACATCCCGGAT 577  
QY 541 GAAGAGTGAAGTCCCTTCGCCCTCAGCGAGCCAGCGTGGGACCACTCTTCCCGG 600  
Db 578 GAAGAGTGAAGTCCCTTCGCCCTCAGCGAGCCAGCGTGGGACCACTCTTCCCGG 637  
QY 601 GAGCAGGCCACGCCCTCTGGGGGCACTTCTCAGGCCAGACAGATTTGATTTCCCGCTGCGG 660  
Db 638 GAGCAGGCCACGCCCTCTGGGGGCACTTCTCAGGCCAGACAGATTTGATTTCCCGCTGCGG 697

```
QY 661 ATCCGTGTCCTCCACCCAGTTTCTTGTCGCATCATCGGAAGAGGGCTTGACCA 715
|||||
Db 698 ATCCGTGTCCTCCACCCAGTTTCTTGTCGCATCATCGGAAGAGGGCTTGACCA 752
|||||

RESULT 7
BM806132 1089 bp mRNA linear EST 05-MAR-2002
LOCUS AGENCOURT_6553922 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:555652
DEFINITION 5', mRNA sequence.
ACCESSION BM806132
VERSION BM806132.1 GI:19122955
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1089)
NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12276 row: c column: 13
High quality sequence stop: 688.
FEATURES
Location/Qualifiers
1..1089
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:555652"
/tissue_type="leiomyosarcoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_71"
/note="Organ: uterus; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 2.1 kb."
ORIGIN
Query Match 20.6%; Score 702; DB 12; Length 1089;
Best Local Similarity 100.0%; Pred. No. 3.3e-267;
Matches 702; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 201 CTACCCCGACCAAGACTGGGCATCGCGCCATCGAGACCCTCTCGGTAAGTGGAATT 260
|||||
Db 1 CTACCCCGACCAAGACTGGGCATCGCGCCATCGAGACCCTCTCGGTAAGTGGAATT 60
|||||

QY 261 GCATGGGAAATCATGGAAGTTGATTACTCAGTCTCTAAAAGCTAAGGACGAGAAAT 320
|||||
Db 61 GCATGGGAAATCATGGAAGTTGATTACTCAGTCTCTAAAAGCTAAGGACGAGAAAT 120
|||||

QY 321 TCAGATTGGAACATCCCTCTCCTCAGTGGAGGTGTGATGACATTTTGGCTCA 380
|||||
Db 121 TCAGATTGGAACATCCCTCTCCTCAGTGGAGGTGTGATGACATTTTGGCTCA 180
|||||

QY 381 ATATGGGACAGTGGAGAATGTGGAACAAGTCAACACACACAGAAACCGCGTGTGCA 440
|||||
Db 181 ATATGGGACAGTGGAGAATGTGGAACAAGTCAACACACACAGAAACCGCGTGTGCA 240
|||||

QY 441 CGTCACATATGCAACAGAGAAAGAGCAAAATAGCCATGAGAGCTAAGCGGATCA 500
|||||
Db 241 CGTCACATATGCAACAGAGAAAGAGCAAAATAGCCATGAGAGCTAAGCGGATCA 300
|||||

QY 501 GTTTGAGAACTACTCTCTCAGATTTCCTACATCCCGATCAAGAGGTGAGTCCCTTC 560
|||||
Db 501 GTTTGAGAACTACTCTCTCAGATTTCCTACATCCCGATCAAGAGGTGAGTCCCTTC 560
|||||

Db 301 GTTTGAGAACTACTCTCTCAGATTTCCTACATCCCGATCAAGAGGTGAGTCCCTTC 360
|||||
QY 561 GCCCCTCAGGAGCCAGCGTGGGACCACTCTTCCCGGAGCAAGCCAGCCCTTGG 620
|||||
Db 361 GCCCCTCAGGAGCCAGCGTGGGACCACTCTTCCCGGAGCAAGCCAGCCCTTGG 420
|||||

QY 621 GGGCACTTCTCAGGCGACACAGATTGATTTCCGCTCGGATCCTGTGTCCTCCACCAAGTT 680
|||||
Db 421 GGGCACTTCTCAGGCGACACAGATTGATTTCCGCTCGGATCCTGTGTCCTCCACCAAGTT 480
|||||

QY 681 TGTGTCCTCATATCGGAAAGAGGGCTTGACCATAAAGAACATCACTAAGCAGACCA 740
|||||
Db 481 TGTGTCCTCATATCGGAAAGAGGGCTTGACCATAAAGAACATCACTAAGCAGACCA 540
|||||

QY 741 GTCCGGGTAGATATCCATAGAAAAGAGAACTCTGGAGCTCGAGAGAGCTGTCCACAT 800
|||||
Db 541 GTCCGGGTAGATATCCATAGAAAAGAGAACTCTGGAGCTCGAGAGAGCTGTCCACAT 600
|||||

QY 801 CCATGCCACCCCGAGAGGGGACTTCTGAAGCATGCGGCATGATTTCTTGAATCATGCAGAA 860
|||||
Db 601 CCATGCCACCCCGAGAGGGGACTTCTGAAGCATGCGGCATGATTTCTTGAATCATGCAGAA 660
|||||

QY 861 AGAGCGAGATGAGACCAAACTAGCCGAGAGAGATTCTCTGAA 902
|||||
Db 661 AGAGCGAGATGAGACCAAACTAGCCGAGAGAGATTCTCTGAA 702
|||||

RESULT 8
BM561242
LOCUS AGENCOURT_6567353 NIH_MGC_88 Homo sapiens cDNA clone IMAGE:5739550
DEFINITION 5', mRNA sequence.
ACCESSION BM561242
VERSION BM561242.1 GI:18906337
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 999)
NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12753 row: a column: 23
High quality sequence stop: 679.
FEATURES
Location/Qualifiers
1..999
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5739550"
/tissue_type="duodenal adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_88"
/note="Organ: small intestine; Vector: pCMV-SPORT6;
Site 1: NotI; Site 2: SalI; Cloned unidirectionally;
oligo-dr primed. Average insert size 1.767 Kb. Library
enriched for full-length clones and constructed by Life
Technologies. Note: this is a NIH_MGC Library."
ORIGIN
Query Match 20.5%; Score 698; DB 12; Length 999;
Best Local Similarity 100.0%; Pred. No. 1.3e-265;
Matches 698; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Qy	1270	ACCACTCCGGATAC	TTCTCAGGCTGTACCCCATCACAGTTTGCCCGCTCCCGCAT	1329
Db	113	ACCACTCCGGATAC	TTCTCAGGCTGTACCCCATCACAGTTTGCCCGCTCCCGCAT	172
Qy	1330	CATCACTCTTAT	CCAGAGCAGGAGTTGTAATCTTCTATCCCAACCCAGGCTGTGGGC	1389
Db	173	CATCACTCTTAT	CCAGAGCAGGAGTTGTAATCTTCTATCCCAACCCAGGCTGTGGGC	232
Qy	1390	GCCATCATCGGGAAG	AGGGGGCACATCAAAACAGCTGCGAGATTCGCCGAGGCCTCT	1449
Db	233	GCCATCATCGGGAAG	AGGGGGCACATCAAAACAGCTGCGAGATTCGCCGAGGCCTCT	292
Qy	1450	ATCBAAGATTGCCCT	TGCGGAGGCCAGAGTCTAGCGAAAGGATGTCATCATCCGGG	1509
Db	293	ATCAAGATTGCCCT	TGCGGAGGCCAGAGTCTAGCGAAAGGATGTCATCATCCGGG	352
Qy	1510	CCACCGGAAGCCCA	GTTCAGAGCCAGGACGGATCTTTCGGAAACTGAAAGAGGAAAC	1569
Db	353	CCACCGGAAGCCCA	GTTCAGAGCCAGGACGGATCTTTCGGAAACTGAAAGAGGAAAC	412
Qy	1570	TTCTTTAAACCCAA	AGGAAGTGAAGCTGGAAAGCGCATATCAGAGTGCCTCTTCCACA	1629
Db	413	TTCTTTAAACCCAA	AGGAAGTGAAGCTGGAAAGCGCATATCAGAGTGCCTCTTCCACA	472
Qy	1630	GCTGGCCGGTGATT	TGCGAAAGTGGCAAGACCGTGAACGAATGCGAGAACTTAAACAGT	1689
Db	473	GCTGGCCGGTGATT	TGCGAAAGTGGCAAGACCGTGAACGAATGCGAGAACTTAAACAGT	532
Qy	1690	GCAGAACTCATCTG	TCCCTGACCAAAACGCCAGATCAAAATGAGGAAGTGATCGTCAGA	1749
Db	533	GCAGAACTCATCTG	TCCCTGACCAAAACGCCAGATCAAAATGAGGAAGTGATCGTCAGA	592
Qy	1750	ATTATCGGGCACT	TTCTTTGTAGCCAGATGCGACAGCGCAAGATCAGGGAAATGTGTACAA	1809
Db	593	ATTATCGGGCACT	TTCTTTGTAGCCAGATGCGACAGCGCAAGATCAGGGAAATGTGTACAA	652
Qy	1810	CAGGTGAAGCAGCAG	GAGCAGAAATACCTTCAGGAGTTCGCTCAGAGCCAGCAAGTGA	1869
Db	653	CAGGTGAAGCAGCAG	GAGCAGAAATACCTTCAGGAGTTCGCTCAGAGCCAGCAAGTGA	712
Qy	1870	GGCTCCCAACAGG	CAACAGGATGAATGTAGCCCTTCCAAACACCTTGACAGAA	1929
Db	713	GGCTCCCAACAGG	CAACAGGATGAATGTAGCCCTTCCAAACACCTTGACAGAA	772
Qy	1930	TGAGACCAACAGC	AGCCAGCCAGATCGGAGCAAAACA	1967
Db	773	TGAGACCAACAGC	AGCCAGCCAGATCGGAGCAAAACA	810

RESULT	9
LOCUS	BM977219
DEFINITION	UI-CF-DUI-ads-n-12-0-UI.s1 linear mRNA EST 21-FEB-2003 UI-CF-DUI-ads-n-12-0-UI.3' mRNA sequence.
ACCESSION	BM977219
VERSION	BM977219.1 GI:19595413
KEYWORDS	EST.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 724) Bonaldo,M.F., Lennon,G. and Soares,M.B. Normalization and subtraction: two approaches to facilitate gene discovery
JOURNAL	Genome Res. 6 (9), 791-806 (1996)
MEDLINE	97044477
PUBMED	8889548
COMMENT	Contact: McCray, PB McCray Lab University of Iowa 2024 University of Iowa Med Labs, Iowa City, IA 52242, USA

```

Tel: 319 356 4866
Fax: 319 356 7171
Email: paul-mccray@uiowa.edu
Tissue Procurement: Dr. M. J. Welsh, University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.resgen.com) or from Open Biosystems
(www.openbiosystems.com) .
Seq primer: M13 FORWARD
POLYA=Yes.
Location/Qualifiers
1. .724
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-CF-DUI-ads-n-12-0-UI"
/tissue_type="Primary Lung Epithelial Cells"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/clone_lib="UI-CF-DUI"
/notes="Organ: Lung; Vector: pT7T3-Pac (Pharmacia) with a
modified polylinker; Site 1: EcoR I; Site 2: Not I;
UI-CF-DUI is a normalized cDNA library containing the
following tissue(s) : Primary lung Epithelial Cells The
library was constructed according to Bonaldo, Lennon and
Soares, Genome Research, 6:791-806, 1996. First strand
cDNA synthesis was primed with an oligo-dT primer
containing a Not I site. Double stranded cDNA was ligated
to an EcoR I adaptor, digested with Not I, and cloned
directionally into pT7T3-Pac vector. The oligonucleotide
used to prime the synthesis of first-strand cDNA contains
a library tag sequence that is located between the Not I
site and the (dT)18 tail. The sequence tag for this
library is GCGTGTAGGC.
TAG TISSUE=Lung Epithelial Cells Tissue nos 359-368
TAG LIB=UI-CF-DUI
TAG_SEQ=GGCTGTAGGC"

```

ORIGIN	Query Match	19.6%	Score 670;	DB 12;	Length 724;
	Best Local Similarity	100.0%;	Pred. No. 2e-254;		
	Matches 670;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	371	TTTTGGCTCAATATGGGACAGTGTGGAACTGTGGAACAAGTCAACACAGACACAGAAACCG	430		
Db	14	TTTTGGCTCAATATGGGACAGTGTGGAACTGTGGAACAAGTCAACACAGACACAGAAACCG	73		
Qy	431	CCGTTGTCAACGTACATATGCAACRAGAGAAAGCAAAAATATAGCCATGGAGAAAGCTAA	490		
Db	74	CCGTTGTCAACGTACATATGCAACRAGAGAAAGCAAAAATATAGCCATGGAGAAAGCTAA	133		
Qy	491	GCGGGCATCAGTTTGAGAACTACTCCTTCAAGATTTCCTACATCCCGATGAAAGAGTGA	550		
Db	134	GCGGGCATCAGTTTGAGAACTACTCCTTCAAGATTTCCTACATCCCGATGAAAGAGTGA	193		
Qy	551	GCTCCCTTCGCCCCCTTCAGCGAGCCGAGCGTGGGGACCACTCTTCCCGGGAGCAAGGCC	610		
Db	194	GCTCCCTTCGCCCCCTTCAGCGAGCCGAGCGTGGGGACCACTCTTCCCGGGAGCAAGGCC	253		
Qy	611	ACGCCCCCTGGGGGACACTTCTCAGCGCAGACAGATTGATTTTCCGCTGCGGATCTCTGGTCC	670		
Db	254	ACGCCCCCTGGGGGACACTTCTCAGCGCAGACAGATTGATTTTCCGCTGCGGATCTCTGGTCC	313		
Qy	671	CCACCCAGTTTGTGGTGCCATCATCGAAAGAGAGGCTTGACCATAAAGAAACATCACTA	730		
Db	314	CCACCCAGTTTGTGGTGCCATCATCGAAAGAGAGGCTTGACCATAAAGAAACATCACTA	373		
Qy	731	AGCAGACCCAGTCCCGGGTAGATATCATAGAAAAAGAGAACTCTGGAGCTGCGAGAAAGC	790		
Db	374	AGCAGACCCAGTCCCGGGTAGATATCATAGAAAAAGAGAACTCTGGAGCTGCGAGAAAGC	433		

```
QY 791 CTGTCACCATCCATCCACCCAGAGGGGCACTTCTGAAGCATGCGCATGATCTTTGAAA 850
Db |||||||
434 CTGTCACCATCCATCCACCCAGAGGGGCACTTCTGAAGCATGCGCATGATCTTTGAAA 493
QY 851 TCATGCAGAAAGAGCAGATGAGACCAAACTAGCGGAAGAGATCTCTGAAAAATCTTTGG 910
Db |||||||
494 TCATGCAGAAAGAGCAGATGAGACCAAACTAGCGGAAGAGATCTCTGAAAAATCTTTGG 553
QY 911 CACCAATGGCTTGGTGGAGACTGATTTGAAAAGAGGAGCAAAATTTGAGAAAATTG 970
Db |||||||
554 CACCAATGGCTTGGTGGAGACTGATTTGAAAAGAGGAGCAAAATTTGAGAAAATTG 613
QY 971 AACATGAACAGGAGCAAGATACAAATCTCATCTTTTCAGAGATTGAGCATATACAACC 1030
Db |||||||
614 AACATGAACAGGAGCAAGATACAAATCTCATCTTTTCAGAGATTGAGCATATACAACC 673
QY 1031 CGAAAGAAC 1040
Db |||||||
674 CGAAAGAAC 683

RESULT 10
BQ212500 836 bp mRNA linear EST 02-MAY-2002
LOCUS AGENCOURT_7675661 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:6095809
5', mRNA sequence.
ACCESSION BQ212500
VERSION BQ212500.1 GI:20392782
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE 1 (bases 1 to 836)
JOURNAL NIH-MGC http://mgi.nci.nih.gov/.
COMMENT National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC/DCTD/DTF
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM13368 row: n column: 02
High quality sequence stop: 670.
FEATURES
source
location/Qualifiers
1..836
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6095809"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 72"
/note="Torgan: skin; Vector: pCMV-SPORT6; Site:1: NotI;
Site:2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 2 kb. Library constructed by Life
Technologies."

ORIGIN
Query Match 18.8%; Score 642; DB 13; Length 836;
Best Local Similarity 100.0%; Pred. No. 2.2e-243;
Matches 642; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2168 GACGATATCCCTTTAGTTGAACCTAACATAGGTGAACGTGTTAAAGCAAGCAAAATGC 2227
Db |||||||
1 GACGATATCCCTTTAGTTGAACCTAACATAGGTGAACGTGTTAAAGCAAGCAAAATGC 60
QY 2228 ACACCCCTTTCTGTGGCAAAATCGTCTGTACATGTGTGTACATATTAGAAAGGAGA 2287
|||
```

```
Db 61 ACACCCCTTTCTGTGGCAAAATCGTCTGTGTACATGTGTGTACATATTAGAAAGGAAGA 120
QY 2288 TGTTAAGATATGTGGCTGTGGGTATACACAGGGTGCCTGACGGGTATATATATTTTAA 2347
Db |||||||
121 TGTTAAGATATGTGGCTGTGGGTATACACAGGGTGCCTGACGGGTATATATATTTTAA 180
QY 2348 ATAATATATCAAAATCAACTCAACTAACTTAACTTAACTTAACTTAACTTAACTTAACTT 2407
Db |||||||
181 ATAATATATCAAAATCAACTCAACTAACTTAACTTAACTTAACTTAACTTAACTTAACTT 240
QY 2408 TTTTAAAGAGAGAAACAGGCTTTTCTAGACTTTTAAAGAAATAAAGTCTTTTGGAGGTCTCAC 2467
Db |||||||
241 TTTTAAAGAGAGAAACAGGCTTTTCTAGACTTTTAAAGAAATAAAGTCTTTTGGAGGTCTCAC 300
QY 2468 GGTGTAGAGAGAGCTTTGAGGCCACCGGCACAAATTCACCCAGAGGGAATCTGTCG 2527
Db |||||||
301 GGTGTAGAGAGAGCTTTGAGGCCACCGGCACAAATTCACCCAGAGGGAATCTGTCG 360
QY 2528 GAAGGACACTCAGCGCAGTCTGGATCACCTGTGTATGTCAACAGAGGGATACCGTCTC 2587
Db |||||||
361 GAAGGACACTCAGCGCAGTCTGGATCACCTGTGTATGTCAACAGAGGGATACCGTCTC 420
QY 2588 CTTGAAGAGAGAAACTCTGTCACTCTCATGCTCTGTAGCTCATACACCCATTTCTCTTT 2647
Db |||||||
421 CTTGAAGAGAGAAACTCTGTCACTCTCATGCTCTGTAGCTCATACACCCATTTCTCTTT 480
QY 2648 GCTTCACAGGTTTAAACTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2707
Db |||||||
481 GCCTTCACAGGTTTAAACTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
QY 2708 TTTATCTCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCC 2767
Db |||||||
541 TTTATCTCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCC 600
QY 2768 ATCCCTCCATCTCAATCCCGTATCTAGCACCCCCCCCCCCCC 2809
Db |||||||
601 ATCCCTCCATCTCAATCCCGTATCTAGCACCCCCCCCCCCCC 642

RESULT 11
BQ674692 647 bp mRNA linear EST 07-OCT-2002
LOCUS UI-CF-DUO-aac-j-06-0-UI.s1 UI-CF-DUO Homo sapiens cDNA clone
DEFINITION UI-CF-DUO-aac-j-06-0-UI 3', mRNA sequence.
ACCESSION BQ674692
VERSION BQ674692.1 GI:23517945
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE 1 (bases 1 to 647)
JOURNAL Bonaldo, M.F., Lennon, G. and Soares, M.B.
MEDLINE Normalization and subtraction: two approaches to facilitate gene
PUBMED discovery
COMMENT Genome Res. 6 (9), 791-806 (1996)
97044477
8989548
Contact: McCray, PB
McCray Lab
University of Iowa
2024 University of Iowa Med Labs, Iowa City, IA 52242, USA
Tel: 319 356 4866
Fax: 319 356 7171
Email: paul-mccray@uiowa.edu
Tissue Procurement: Dr. M. J. Welsh, University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.resgen.com) or from Open Biosystems
(www.openbiosystems.com).
Seq primer: M13 FORWARD
```

FEATURES  
source POLYA=Yes Location/Qualifiers  
1. .647  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="UI-CF-DU0-aac-j-06-0-UI"  
/tissue\_type="Primary Lung Epithelial Cells"  
/dev\_stage="Adult"  
/lab\_host="DH10B (Life Technologies) (T1 phage resistant)"  
/clone\_lib="UI-CF-DU0"  
/notes="Organ: Lung; Vector: pT7T3-Pac (Pharmacia) with a modified polylinker; Site\_1: EcoR I; Site\_2: Not I; UI-CF-DU0 is a cDNA library containing the following tissue(s): Primary lung Epithelial Cells The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT7T3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is GGCTGTAGGC.  
TAG\_TISSUE=Lung Epithelial Cells Tissue nos 359-368  
TAG\_LIB=UI-CF-DU0  
TAG\_SEQ=GGCTGTAGGC"

ORIGIN  
Query Match 18.6%; Score 634; DB 13; Length 647;  
Best Local Similarity 100.0%; Pred. No. 3.8e-240;  
Matches 634; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 371 TTTTGGCTCAATATGGACAGTGGAGATCTGGAACAAGTCAACACAGACACAGAAACCG 430  
DB 14 TTTTGGCTCAATATGGACAGTGGAGATCTGGAACAAGTCAACACAGACACAGAAACCG 73  
QY 431 CCGTTGTCAACGTCACATATGCAACAGAGAAGCAAAAATAGCCATCGGAAGCTAA 490  
DB 74 CCGTTGTCAACGTCACATATGCAACAGAGAAGCAAAAATAGCCATCGGAAGCTAA 133  
QY 491 GCGGGCATAGTTTGAGAACTACTCTTCAAGATTTCTACATCCCGGATGAGAGGTGA 550  
DB 134 GCGGGCATAGTTTGAGAACTACTCTTCAAGATTTCTACATCCCGGATGAGAGGTGA 193  
QY 551 GCTCCCTTCGCCCTTCAGCGAGCCAGCGTGGGGACCACTCTTCCCGGGAGCAAGGCC 610  
DB 194 GCTCCCTTCGCCCTTCAGCGAGCCAGCGTGGGGACCACTCTTCCCGGGAGCAAGGCC 253  
QY 611 ACGCCCTTGGGGGACCTTCTCAGGCGACAGAGATTGATTTCCCGCTCGGATCCTGGTCC 670  
DB 254 ACGCCCTTGGGGGACCTTCTCAGGCGACAGAGATTGATTTCCCGCTCGGATCCTGGTCC 313  
QY 671 CCACCCAGTTTGTGTGTCATCATCGGAAGAGGGCTTGACCATAAAGACATCACTA 730  
DB 314 CCACCCAGTTTGTGTGTCATCATCGGAAGAGGGCTTGACCATAAAGACATCACTA 373  
QY 731 AGCAGACCCAGTCCCGGTAGATATCCATAGAAAGAGAACTCTGGAGCTGCGAGAGGC 790  
DB 374 AGCAGACCCAGTCCCGGTAGATATCCATAGAAAGAGAACTCTGGAGCTGCGAGAGGC 433  
QY 791 CTGTACCATTCATGCCACCCGAGGGGACTTCTGAAGCATGCGCATGATTTTGAAA 850  
DB 434 CTGTACCATTCATGCCACCCGAGGGGACTTCTGAAGCATGCGCATGATTTTGAAA 493  
QY 851 TCATGAGAGAGGCGAGATGAGCAAACTAGCCGAGAGATTCCTCTGAATCTTGG 910  
DB 494 TCATGAGAGAGGCGAGATGAGCAAACTAGCCGAGAGATTCCTCTGAATCTTGG 553  
QY 911 CACCAATGGCTTGGTGGAGACTGATTGGAAAGAGGAGAAATTCGAAGAAATTG 970  
DB 554 CACCAATGGCTTGGTGGAGACTGATTGGAAAGAGGAGAAATTCGAAGAAATTG 613

QY 971 AACATGAACACGAGCAACAGATAAACAATCTCATC 1004  
DB 614 AACATGAACACGAGCAACAGATAAACAATCTCATC 647

RESULT 12  
BG748346 889 bp mRNA linear EST 15-MAY-2001  
LOCUS 602705902F1 NIH\_MGC\_43 Homo sapiens cDNA clone IMAGE:4842553 5',  
DEFINITION mRNA sequence.  
ACCESSION BG748346  
VERSION BG748346.1 GI:14058999  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 889)  
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-r@mail.nih.gov  
Tissue Procurement: ATCC  
CDNA Library Preparation: Ling Hong/Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLCMI677 row: c column: 02  
High quality sequence stop: 773.

FEATURES  
source Location/Qualifiers  
1. .889  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4842553"  
/tissue\_type="normal pigmented retinal epithelium"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_43"  
/note="Organ: eye; Vector: pOTB7; Site\_1: XhoI; Site\_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH\_MGC Library. |"

ORIGIN  
Query Match 18.1%; Score 617; DB 12; Length 889;  
Best Local Similarity 99.7%; Pred. No. 1.7e-233;  
Matches 717; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 72 GATGAACAAGCTTTACATCGGGAACCTGAGCCCGCGTCACCGCAGACCTCCGGCA 131  
DB 2 GATGAACAAGCTTTACATCGGGAACCTGAGCCCGCGTCACCGCAGACCTCCGGCA 61  
QY 132 GCTCTTTGGGACAGAAAGCTGCCCTGGGGACAGAGTCTCTGAAGTCCGGCTACGC 191  
DB 62 GCTCTTTGGGACAGAAAGCTGCCCTGGGGACAGAGTCTCTGAAGTCCGGCTACGC 121  
QY 192 CTTCTGTGGACTACCCGACACAGAACTGGCGCATCGGCATCGAGACCTCTCGGGTAA 251  
DB 122 CTTCTGTGGACTACCCGACACAGAACTGGCGCATCGAGACCTCTCGGGTAA 191  
QY 252 AGTGGAAATTCATGGGAAATCATGGAAGTTGATTACTCAGTCTCTAAAAGCTAAGGAG 311  
DB 182 AGTGGAAATTCATGGGAAATCATGGAAGTTGATTACTCAGTCTCTAAAAGCTAAGGAG 241  
QY 312 CAGGAAATTCAGATTCGAAACATCCCTCTCACTTCAGTGGGAGGTGTGGATGACT 371

Db 242 CAGGAAATTCAGATTCCCTCCTCAGTGGAGGTGTTGGATGACT 301  
Qy 372 TTTGGCTCAATATGGGACGAGTGGGAAATGTGGAACTACACAGACACAGAAACCGC 431  
Db 302 TTTGGCTCAATATGGGACGAGTGGGAAATGTGGAACTACACAGACACAGAAACCGC 361  
Qy 432 CGTTGTCAACGTACATATGCAACAAAGAGAGAAACAAATATAGCATGGAGAACTAAG 491  
Db 362 CGTTGTCAACGTACATATGCAACAAAGAGAGAAACAAATATAGCATGGAGAACTAAG 421  
Qy 492 CGGGCATCAGTTTGAGAACTACTCTTCAAGATTTTCTTACATCCCGGATGAAGAGGTGAG 551  
Db 422 CGGGCATCAGTTTGAGAACTACTCTTCAAGATTTTCTTACATCCCGGATGAAGAGGTGAG 481  
Qy 552 CTCCTCTTCCGCTCAGGAGCCAGGTTGGGACCACTCTTCCCGGAGCAAGGCCA 611  
Db 482 CTCCTCTTCCGCTCAGGAGCCAGGTTGGGACCACTCTTCCCGGAGCAAGGCCA 541  
Qy 612 CGCCCTGGGGGCACTTCTCAGGCGCAGACAGATTGATTTCCGCTGCGGATCCTGGTCCC 671  
Db 542 CGCCCTGGGGGCACTTCTCAGGCGCAGACAGATTGATTTCCGCTGCGGATCCTGGTCCC 601  
Qy 672 CACCCAGTTTGTGGTGCATCATCGGAAAGAGGGCTTGACCAATAAAGAAACATCACTAA 731  
Db 602 CAACCCAGTTTGTGGTGCATCATCGGAAAGAGGGCTTGACCAATAAAGAAACATCACTAA 661  
Qy 732 GCAGACCCAGTCCCGGGTAGATATCCATAGAAAGAGAACTCTGGAGCTGCAGAGAAGC 790  
Db 662 GCAGACCCAGTCCCGGGTAGATATCCATAGAAAGAGAACTCTGGAGCTGCAGAGAAGC 720

RESULT 13  
BG575889 781 bp mRNA linear EST 10-APR-2001  
LOCUS 602598315F1 NIH\_MGC\_87 Homo sapiens cDNA clone IMAGE:4707243 5',  
DEFINITION mRNA sequence.  
ACCESSION BG575889  
VERSION BG575889.1 GI:13583542  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 781)  
NIH-MGC <http://mgc.nci.nih.gov/>.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
Tissue Procurement: DCTD/DTF  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: L1AM10576 row: m column: 04  
High quality sequence stop: 773.  
Location/Qualifiers  
1..781  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4707243"  
/tissue\_type="mammary adenocarcinoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_87"  
/note="Organ: breast; Vector: pCMV-SPORT6; Site 1: NotI;  
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.  
Average insert size 1.383 kb. Library enriched for  
full-length clones and constructed by Life Technologies.  
Note: this is a NIH\_MGC Library."

FEATURES  
source

ORIGIN  
Query Match 17.4%; Score 594; DB 12; Length 781;  
Best Local Similarity 99.7%; Pred. No. 2.3e-224;  
Matches 694; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
Qy 1845 AGTCGCTTCACAGCGCAGCAAGTGGAGTCCACAGGACACAGCAAAACACCGATGAAT 1904  
Db 25 AGTCGCTTCACAGCGCAGCAAGTGGAGTCCACAGGACACAGCAAAACACCGATGAAT 84  
Qy 1905 GTAGCCCTTCCAAACCTTCAGCAATGAGACAAACGAGCCAGCAGATCGGAGCAAA 1964  
Db 85 GTAGTCTCTTCCAAACCTTCAGCAATGAGACAAACGAGCCAGCAGATCGGAGCAAA 144  
Qy 1965 CCAAGACCATCTGAGGATCAGAACTCTGGAGGGGGCCAGGGACTCTCCGAGGGCC 2024  
Db 145 CCAAGACCATCTGAGGATCAGAACTCTGGAGGGGGCCAGGGACTCTCCGAGGGCC 204  
Qy 2025 TGAGAACCCCAAGGGCCGAGGAGGGGGGGAAGGTACAGCAGGTTTCCAGAAACCCAG 2084  
Db 205 TGAGAACCCCAAGGGCCGAGGAGGGGGGGAAGGTACAGCAGGTTTCCAGAAACCCAG 264  
Qy 2085 AGCCCGGCTCTCCGCCCCCAGGGCTTCTGCAAGCTTCAGGCATCCACTTCCACATCCAC 2144  
Db 265 AGCCCGGCTCTCCGCCCCCAGGGCTTCTGCAAGCTTCAGGCATCCACTTCCACATCCAC 324  
Qy 2145 TCGGATCTCTCTGAACTCCCAAGCGCTATCCCTTTTCTGTTGGCAAAATCGTCTCTGTACATGT 2204  
Db 325 TCGGATCTCTCTGAACTCCCAAGCGCTATCCCTTTTCTGTTGGCAAAATCGTCTCTGTACATGT 384  
Qy 2205 GTGTTTCAAAGCCCAAGCAAAATGCACACCCCTTTTCTGTTGGCAAAATCGTCTCTGTACATGT 2264  
Db 385 GTGTTTCAAAGCCCAAGCAAAATGCACACCCCTTTTCTGTTGGCAAAATCGTCTCTGTACATGT 444  
Qy 2265 GTGTACATATTAGAAAGGAAAGATGTTAAAGATATGTGGCTGTGGTTTACACAGGGTCC 2324  
Db 445 GTGTACATATTAGAAAGGAAAGATGTTAAAGATATGTGGCTGTGGTTTACACAGGGTCC 504  
Qy 2325 TGCAGCGGTAAATATATTAGAAATATATATCAAAATCACTCAACTCTCAATTTTAA 2384  
Db 505 TGCAGCGGTAAATATATTAGAAATATATATCAAAATCACTCAACTCTCAATTTTAA 564  
Qy 2385 ATCAATATTAAATTTTTTTTCTTTTAAAGAGAAAGCAGGCTTTCTAGACTTTTAAAGA 2444  
Db 565 ATCAATATTAAATTTTTTTTCTTTTAAAGAGAAAGCAGGCTTTCTAGACTTTTAAAGA 624  
Qy 2445 ATAAAGTCTTTGGAGGTCTCACGGTGTAGAGAGAGCTTTGAGGCCACCCGCAAAAT 2504  
Db 625 ATACAGTCTTTGGAGGTCTCACGGTGTAGAGAGAGCTTTGAGGCCACCCGCAAAAT 684  
Qy 2505 TCACCCAGAGGGAATCTCGTCGGAAGGACACTCAC 2540  
Db 685 TCACCCAGAGGGAATCTCGTCGGAAGGACACTCAC 720  
RESULT 14  
BG575825  
LOCUS 7965128 NIH\_MGC\_72 Homo sapiens cDNA clone IMAGE:6162660  
DEFINITION 5', mRNA sequence.  
ACCESSION BG575825  
VERSION BG575825.1 GI:22669777  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 894)  
NIH-MGC <http://mgc.nci.nih.gov/>.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)







GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: July 16, 2004, 10:30:58 ; Search time 149 Seconds  
(without alignments)  
12940.296 Million cell updates/sec

Title: US-09-270-437D-6

Perfect score: 6121

Sequence: 1 ggcagcgaggagggcgagga.....aaccttgaaatgtttattt 3412

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 3172214

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODE=frame+ n2p.model -DEV=xlp  
-Q-/cgn2.1/USPTO.spool.p/US09270437/runat.16072004.113048.13308/app.query.fasta\_1.3591  
-DB=A Geneseq 29Jan04 -QFMT=fastan -SUFFIX=rag -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS-human40.cdi  
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09270437@cgn.1.253@runat.16072004.113048.13308 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : A Geneseq 29Jan04:\*  
1: Geneseqp1980s:\*  
2: Geneseqp1990s:\*  
3: Geneseqp2000s:\*  
4: Geneseqp2001s:\*  
5: Geneseqp2002s:\*  
6: Geneseqp2003as:\*  
7: Geneseqp2003bs:\*  
8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3094	50.5	614	4 ABG06794	Abg06794 Novel hum
2	3065.5	50.1	620	4 AAU16163	Aau16163 Human nov
3	3065.5	50.1	620	6 ABU55232	Abu55232 Human nov
4	2917.5	47.7	594	4 ABG06795	Abg06795 Novel hum
5	2806.5	45.9	556	5 ABG96346	Abg96346 Human ova
6	2553	41.7	555	6 ABU89799	Aay89799 Novel hum
7	2019.5	33.0	577	2 AAY30649	Aay30649 A murine
8	1957.5	32.0	579	3 AAB11365	Aab11365 Human lun
9	1957.5	32.0	579	5 ABB75053	Abb75053 Human lun
10	1957.5	32.0	579	5 ABB74997	Abb74997 Human lun

11	1957.5	32.0	579	5 ABB75054	Abb75054 Human lun
12	1957.5	32.0	579	5 ABB75054	Abb75054 Human lun
13	1957.5	32.0	579	5 ABB75054	Abb75054 Human lun
14	1957.5	32.0	579	5 ABB75054	Abb75054 Human lun
15	1957.5	32.0	579	7 ADA28536	Ada28536 Recombina
16	1957.5	32.0	579	7 ADA28539	Ada28539 Recombina
17	1957.5	32.0	579	7 ADA28438	Ada28438 Human lun
18	1952.5	31.9	586	5 ABB75048	Abb75048 Human lun
19	1952.5	31.9	586	5 ABB75048	Abb75048 Human lun
20	1952.5	31.9	586	7 ADA28517	Ada28517 Recombina
21	1950.5	31.9	579	3 AAB11328	Aab11328 Human lun
22	1950.5	31.9	579	5 ABB74960	Abb74960 Human lun
23	1950.5	31.9	579	5 ABB74960	Abb74960 Human lun
24	1950.5	31.9	579	7 ADA28286	Ada28286 Human lun
25	1950.5	31.9	579	7 ADA28286	Ada28286 Human lun
26	1950.5	31.9	579	7 ADE53471	Ades3471 Human src
27	1877.5	30.7	619	4 ABG21963	Abg21963 Novel hum
28	1804.5	29.5	583	4 ABG12592	Abg12592 Novel hum
29	1161	19.0	319	4 AAM93826	Aam93826 Human pol
30	910	14.9	192	4 AAU16164	Aau16164 Human nov
31	910	14.9	192	6 ABU55233	Abu55233 Human nov
32	898	14.7	261	4 AAU16161	Aau16161 Human nov
33	898	14.7	261	6 ABU55230	Abu55230 Human nov
34	868	14.2	250	4 AAU16579	Aau16579 Human nov
35	868	14.2	250	6 ABU55648	Abu55648 Human nov
36	865.5	14.1	558	4 ABB58367	Abb58367 Drosophil
37	840.5	13.7	187	6 AAO23971	Aao23971 Human IGF
38	803	13.1	171	4 AAU16166	Aau16166 Human nov
39	803	13.1	171	6 ABU55235	Abu55235 Human nov
40	750	12.3	171	4 AAU16583	Aau16583 Human nov
41	750	12.3	171	6 ABU55652	Abu55652 Human nov
42	501	8.2	97	3 AAG03261	Aag03261 Human sec
43	458	7.5	148	4 ABG21962	Abg21962 Novel hum
44	452	7.4	266	4 ABG12593	Abg12593 Novel hum
45	448	7.3	209	4 ABG21961	Abg21961 Novel hum

## ALIGNMENTS

RESULT 1  
ABG06794  
ID ABG06794 standard; protein; 614 AA.  
AC ABG06794;  
XX  
XX  
DT 13-FEB-2002 (first entry)  
XX  
DE Novel human diagnostic protein #6785.  
XX  
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
KW food supplement; medical imaging; diagnostic; genetic disorder.  
XX  
OS Homo sapiens.  
XX  
PN WC200175067-A2.  
XX  
PD 11-OCT-2001.  
XX  
PF 30-MAR-2001; 2001WO-US008631.  
XX  
PR 31-MAR-2000; 2000US-00540217.  
PR 23-AUG-2000; 2000US-00649167.  
XX  
(HYSE-) HYSEQ INC.  
PA Drmanac RT, Liu C, Tang YT;  
PI  
XX  
XX  
DR WPI; 2001-639362/73.  
DR N-PSDB; AAS70981.  
XX  
XX  
PT New isolated polynucleotide and encoded polypeptides, useful in  
PT diagnostics, forensics, gene mapping, identification of mutations  
PT responsible for genetic disorders or other traits and to assess

PT biodiversity.  
 PS Claim 20; SEQ ID NO 37153; 103pp; English.  
 XX The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (II) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic amino acid sequences of the invention. Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

XX SQ Sequence 614 AA;  
 Alignment Scores:  
 Pred. No.: 5,53e-271 Length: 614  
 Score: 3094.00 Matches: 610  
 Percent Similarity: 99.03% Conservative: 0  
 Best Local Similarity: 99.03% Mismatches: 4  
 Query Match: 50.55% Indels: 2  
 DB: 4 Gaps: 0

US-09-270-437D-6 (1-3412) x ABG06794 (1-614)

QY 23 GCCGGGTACCGGGCGGGGAGCGCGGCTCTCGGGAGAGACGGATGATGAACAGC 82  
 Db 1 AlaGlyTyrArgArg-ArgGluProArgPro-GlnGlyLysArgArgMetMetAsnLysL 20

QY 83 TTTACATCGGAACCTGAGCCCCCGTACCGCGAGACCTCCGGCAGCTCTTTGGGG 142  
 Db 20 euTyrIleGlyAsnLeuSerProAlaValThrAlaAspLeuArgGlnLeuPheGlyA 40

QY 143 ACAGGAAGCTCCCTCGCGGAGCAGGTCTGCTGAAGTCCGGCTACGCTTCGTGACT 202  
 Db 40 spArgLysLeuProLeuAlaGlyGlnValLeuLysSerGlyTyrAlaPheValAspT 60

QY 203 ACCCGGACCCAGAACCTGGGCCATCCGGCCATCGAGACCTCTCGGGTAAAGTGAATTGC 262  
 Db 60 YrProAspGlnAsnTrpAlaIleArgAlaIleGluThrLeuSerGlyLysValGluLeuH 80

QY 263 ATGGGAAATCATGGAATGATTACTCAGTCTCTAAAGCTAAGGACGAGGAAATTC 322  
 Db 80 isGlyLysIleMetGluValAspTyrSerValSerLysLysLeuArgSerArgLysIleG 100

QY 323 AGATTGGAACATCCCTCCCTCACCTGCAGTGGAGGTGTTGGATGGACTTTGGCTCAAT 382  
 Db 100 InIleArgAsnIleProProHisLeuGlnTrpGluValLeuAspGlyLeuLeuAlaGlnT 120

QY 383 ATGGGACAGTGGAGAATGTGAACAAGTCAACACAGACACAGAAACCGCGCTTGTCAACG 442  
 Db 120 YrGlyThrValGluAsnValGluGlnValAsnThrAspThrGluThrAlaValValAsnV 140

QY 443 TCACATATGCAACAGAGAGAGCAAAATAGCCATGGAGACCTTAAGCGGGCATCAGT 502  
 Db 140 alThrTyrAlaThrArgGluGluAlaLysIleAlaMetGluLysLeuSerGlyHisGlnP 160

QY 503 TTGAGAATCTCTCTCAAGATTTCTACATCCCGGATGAAGAGGTGAGCTCCCTTCGC 562  
 Db 160 heGluAsnTyrSerPheLysIleSerTyrIleProAspGluGluValSerSerProSerP 180

QY 563 CCCCTCAGCAGCCAGCGTGGGAGCCACTCTTCCCGGAGAGAGGCCACGCCCTCGGG 622  
 Db 180 roProGlnArgAlaGlnArgGlyAspHisSerSerArgGlnGlnGlyHisAlaProGlyG 200

QY 623 GCATCTTCCAGCCAGACAGATTGATTTCCCGCTCGGATCTGTCTCCACCCCAAGTTTG 682  
 Db 200 lyThrSerGlnAlaArgGlnIleAspPheProLeuArgIleLeuValProThrGlnPheV 220

QY 683 TTGTGGCATCATCGAAAGAGGGCTTGACCATTAAGAAACATCACTAAGCAGACCCAGT 742  
 Db 220 alGlyAlaIleIleGlyLysGluGlyLeuThrIleLysAsnIleThrLysGlnThrGlnS 240

QY 743 CCCGGGTAGATATCCATAGAAAGACACTCTGAGCTGCAGAGAGCCTCTCCACATCC 802  
 Db 240 erArgValAspIleHisArgLysGluAsnSerGlyAlaAlaGluLysProValThrIleH 260

QY 803 ATGCACCCAGAGGGGACTTCTGAAGCATCCCGCATGATTCTTGAATCATGAGAGAAG 862  
 Db 260 isAlaThrProGluGlyThrSerGluAlaCysArgMetIleLeuGluIleMetGlnLysG 280

QY 863 AGCAGATGAGACCAACTAGCCGAGAGATTCTCTGAAATCTTGGCACACATGCTGCT 922  
 Db 280 luAlaAspGluThrLysLeuAlaGluGluIleProLeuLysIleLeuAlaHisAsnGlyL 300

QY 923 TGGTTGGAAGACTGATTGGAAAGAGGACAGAAATTTGAAGAAATTTGAACATGAACAG 982  
 Db 300 euValGlyArgLeuIleGlyLysGluArgAsnLeuLysIleGluHisGluThrG 320

QY 983 GGACCAAGATAACATCTCATCTTTGAGGATTTGACATATACAACCCGGAAGAACCA 1042  
 Db 320 lyThrLysIleThrIleSerSerLeuGlnAspLeuSerIleTyrAsnProGluArgThrI 340

QY 1043 TCATCTGAGGGCAGACAGTTGAGCCCTGTGCGCATGTGTGAGATGAGATATGAAGAAGC 1102  
 Db 340 leThrValLysGlyThrValGluAlaCysAlaSerAlaGluIleGluIleMetLysLysL 360

QY 1103 TGCTGTAGGCTTTGAAATGATATGCTGGCTGTAAACCAACAGCAATCTGATCCAG 1162  
 Db 360 euArgGluAlaPheGluAsnAspMetLeuAlaValAsnGlnGluAlaAsnLeuIleProG 380

QY 1163 GGTTGAACCTCAGCGCATCTGGCATCTTTTCAACAGGACTGTCCGTGCTATCTCCACAG 1222  
 Db 380 lyLeuAsnLeuSerAlaLeuGlyIlePheSerThrGlyLeuSerValLeuSerProProA 400

QY 1223 CAGGCCCCCGGAGCTCCCCCGCTGCCCTACACCCCTTACCTACCCACTCCCGAT 1282  
 Db 400 laGlyProArgGlyAlaProProAlaAlaProTyrHisProPheThrThrHisSerGlyT 420

QY 1283 ACTTCTCAGCTGTACCCCATCACCAGTTTGGCCCGTTCGCCCATCATCATCTTATC 1342  
 Db 420 yrPheSerSerLeuTyrProHisGlnPheGlyProPheProHisHisSerTyrP 440

QY 1343 CAGAGCAGGAGATGTGAATCTTTCATCCCAACCCAGGCTGTGGCGCCATCATCGGA 1402  
 Db 440 roGluGlnGluIleValAsnLeuPheIleProThrGlnAlaValGlyAlaIleIleGlyL 460

QY 1403 AGAAGGGGACACATCAACAGCTGGGAGATTCGGCGGAGCCTCTCATAGATTCGCC 1462  
 Db 460 yLysGlyAlaHisIleLysGlnLeuAlaArgPheAlaGlyAlaSerIleLysIleAlaP 480

QY 1463 CTGGGGAAGGCCACAGCTCAGCAAGAGATGTGTCATCATCCCGGCCACCCGGAAGCCC 1522  
 Db 480 roAlaGluGlyProAspValSerGluArgMetValIleIleThrGlyProProGluAlaG 500

QY 1523 AGTTCAAGGCCAGGAGCGGATCTTTGGGAACTGAAAGAGGAACTCTTTTAAACCCCA 1582  
 Db 500 InPheLysAlaGlnGlyArgIlePheGlyLysLeuLysGluGluAsnPhePheAsnProL 520

QY 1583 AAGAAGAGTGAAGCTGGAAGCGCATATCAGAGTGCCTCTCCACAGCTGGCGGCTGA 1642  
 Db 520 ysGluGluValLysLeuGluAlaHisIleArgValProSerSerThrAlaGlyArgValI 540





```
QY 973 CATGAACAGGACCAAGATAAACAATCTCATCTTTGACGAGTTTGACATATACAACCCG 1032
Db |||||||
QY 323 HisGluThrGlyThrLysIleThrIleSerSerLeuGlnAspLeuSerIleTyrAsnPro 342
Db |||||||
QY 1033 GAAAGAACCATCACTGTGAAGGGCACAGTTGAGCGCTGTGCGAGTGTGAGATAGAGATT 1092
Db |||||||
QY 343 GluArgThrIleThrValLysGlyThrValGluAlaCysAlaSerAlaGluIleGluIle 362
QY 1093 ATGAAGAAGCTGCTGAGGCGCTTTGAAATGATATGCTGGCTGTAAACCAACAGCCAAAT 1152
Db |||||||
QY 363 MetLysLysLeuArgGluAlaPheGluAsnAspMetLeuAlaValAsnGlnGlnAlaAsn 382
QY 1153 CTGATCCACAGGCTTGAACCTCAGCGCACTTGGCATCTTTCAACAGGACTGTCCTGCTA 1212
Db |||||||
QY 383 LeuIleProGlyLeuAsnLeuSerAlaLeuGlyIlePheSerThrGlyLeuSerValLeu 402
QY 1213 TCTCCACAGCAGGCGCCCGGAGCTCCCGCCCTGCCCTACCAACCCCTTCCATACC 1272
Db |||||||
QY 403 SerProProAlaGlyProArgGlyAlaProProAlaAlaProTyrHisProPheThrThr 422
QY 1273 CACTCCGGATCTCTCCAGCTGTACCCCATCACAGTTTGGCCCGTTCGCCGATCAT 1332
Db |||||||
QY 423 HisSerGlyTyrPheSerSerLeuTyrProHisGlnPheGlyProPheProHisHis 442
QY 1333 CACTCTTATCCAGACGAGGAGATTGTGAATCTCTTCATCCCAACCCAGGCTGTGGCGCC 1392
Db |||||||
QY 443 HisSerTyrProGluGlnGluIleValAsnLeuPheIleProThrGlnAlaValGlyAla 462
QY 1393 ATCATCGGGAAGAGGGGGCACACATCAACAGCTGGCGAGATTCCGCGGAGCTCTATC 1452
Db |||||||
QY 463 IleIleGlyLysGlyAlaHisIleLysGlnLeuAlaArgPheAlaGlyAlaSerIle 482
QY 1453 AGATTCCCTCCGGAAGCCACAGCTCAGCGAAAGGATGTCATCATCACCGGCCCA 1512
Db |||||||
QY 483 LysIleAlaProAlaGluGlyProAspValSerGluArgMetValIleIleThrGlyPro 502
QY 1513 CCGAAGCCAGTTTCAAGGCCCGGAGCGGATCTTTGGGAACTGAAAGAGGAAACTTC 1572
Db |||||||
QY 503 ProGluAlaGlnPheLysAlaGlnGlyArgIlePheGlyLysLeuLysGluGluAsnPhe 522
QY 1573 TTTAACCCCAAGAAGTGAAGCTGGAAGCGCATATCAGAGTGCCTCTTCCACAGCT 1632
Db |||||||
QY 523 PheAsnProLysGluGluValLysLeuGlu**HisIleArgValProSerSerThrAla 542
QY 1633 GCGCGGTGATTGGCAAGTGGCAAGACCGTGAACGAATCGCAAACTTAACAGGCA 1692
Db |||||||
QY 543 GlyArgValIleGlyLysGlyLysThrValAsnGluLeuGlnAsnLeuThrSerAla 562
QY 1693 GAAGTTCATCTGCTCTGACCAAAACGCCAGATGAAATCAGGAAGTGATCGTCAGATT 1752
Db |||||||
QY 563 GluValIleValProArgAspGlnThrProAspGluAsnGluGluValIleValArgIle 582
QY 1753 ATCGGCACTCTTTGTACCCAGACTGCAGCGCAAGATCAGGGAATTTGACACAG 1812
Db |||||||
QY 583 IleGlyHisPheAlaSerGlnThrAlaGlnArgLysIleArgGluIleValGlnGln 602
QY 1813 GTGAAGCAGCAGGAGCAGAATACCTTCAGGAGTGCCTCAGCGCAGCAGAG 1866
Db |||||||
QY 603 ValLysGlnGlnGluGlnLysTyrProGlnGlyValAlaSerGlnArgSerLys 620
RESULT 3
ABU55232
ID ABU55232 standard; protein; 620 AA.
XX AC ABU55232;
XX DT 18-MAR-2003 (first entry)
XX DE Human novel polypeptide #319.
XX KW Human; neural disorder; immune system disorder; renal disorder;
KW muscular disorder; respiratory disease; reproductive disorder;
```

```
KW gastrointestinal disorder; pulmonary disorder; cardiovascular disorder;
KW hyperproliferative disorder; inflammatory disease; allergic reaction;
KW blood related disorder; cancer; immunosuppressive; antiinflammatory;
KW cardiovascular; nephrotropic; cytostatic; antiallergic; thrombolytic;
KW haemostatic; antiarteriosclerotic.
XX Homo sapiens.
OS
XX US2002132753-A1.
PN 19-SEP-2002.
XX 17-JAN-2001; 2001US-00764864.
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 28-JUN-2000; 2000US-0214886P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 22-AUG-2000; 2000US-0226868P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 08-SEP-2000; 2000US-0231413P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 27-SEP-2000; 2000US-0235834P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 23-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239335P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241809P.
PR 01-NOV-2000; 2000US-0244617P.
PR 17-NOV-2000; 2000US-0249299P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
XX (ROSE/) ROSEN C A.
PA (RUBE/) RUBEN S M.
PA (BARA/) BARASH S C.
XX Rosen CA, Ruben SM, Barash SC;
PI WPI; 2003-147444/14.
XX N-PSDB; ABX73491.
DR
```

XX New polypeptides and nucleic acids, useful in gene therapy for treating,  
PT inhibiting or preventing e.g. neural, immune system, muscular,  
PT respiratory, reproductive, gastrointestinal, pulmonary, cardiovascular or  
PT renal disorders.  
XX Claim 11; SEQ ID NO 1116; 402pp; English.

XX The invention relates to human novel polypeptides and their associated  
CC polynucleotides. The polypeptides and polynucleotides are useful in gene  
CC therapy for treating, inhibiting or preventing neural disorders, immune  
CC system disorders (e.g. systemic lupus erythematosus, rheumatoid arthritis  
CC and multiple sclerosis), muscular disorders, respiratory diseases (e.g.  
CC nasal vestibulitis, nasal polyps and sinusitis), reproductive disorders,  
CC gastrointestinal disorders, pulmonary disorders, cardiovascular disorders  
CC (e.g. congenital heart defects, Ebstein's anomaly and hypoplastic left  
CC heart syndrome), renal disorders (e.g. acute kidney failure and end-stage  
CC renal disease), hyperproliferative disorders (e.g. Hodgkin's disease and  
CC leukaemia), inflammatory diseases (e.g. septic shock, bursitis and  
CC appendicitis), allergic reactions and conditions (e.g. asthma), blood  
CC related disorders (e.g. thrombosis, atherosclerosis and myocardial  
CC infarction) and cancerous diseases. Sequences ABUS4914-ABUS5699 and  
CC ABUS5748 represent human novel polypeptides of the invention  
XX

SQ Sequence 620 AA;

Alignment Scores:  
Pred. No.: 2,146-268 Length: 620  
Score: 3065.50 Matches: 603  
Percent Similarity: 97.73% Conservative: 1  
Best Local Similarity: 97.57% Mismatches: 11  
Query Match: 50.08% Indels: 3  
DB: 6 Gaps: 1

US-09-270-437D-6 (1-3412) x ABUS5232 (1-620)

QY 22 CGCGGGTACCGCGCGCGGGAGCGCGGCTCTCGGGAGAGACGCG-----ATG 72  
DB 3 ArgArgTyrAlaCysArgTyrArgSerGlyIleProGlySerThrHisAlaSerGlyMet 22  
QY 73 ATGAACAAGCTTATCGGNAACCTGAGCCCGCGGCTACCGCGGACGACCTCGGCGAG 132  
DB 23 MetAsnLysLeuTyrIleGlyAsnLeuSerProAlaValThrAlaAspLeuArgGln 42  
QY 133 CTCCTTTGGGGCAGGAAGTCCCTCGCGGACAGGTCCTGCTCAAGTCCGGCTACGCC 192  
DB 43 LeuPheGlyAspArgLysLeuProLeuAlaGlyGlnValLeuLysSerGlyTyrAla 62  
QY 193 TTCGTGGACTACCCGACAGAACTGGGCCATCGCGCCATCGAGACCTCTCGGGTAAA 252  
DB 63 PheValAspTyrProAspGlnAsnTrpAlaIleArgAlaIleGluThrLeuSerGlyLys 82  
QY 253 GTGGAAATTCAGATTCGAAACATCTGGAAGTTGATTTACTCAGTCTCTAAAAGCTAAGGAC 312  
DB 83 ValGluLeuHisGlyLysIleMetGluValAspTyrSerValSerLysLysLeuArgSer 102  
QY 313 AGGAAATTCAGATTCGAAACATCTGCTCACTGCGAGTGGAGGTCTTGGATGCACTT 372  
DB 103 ArgLysIleGlnIleArgAsnIleProProHisLeuGlnTrpGluValLeuAspGlyLeu 122  
QY 373 TTGGTCAATATGGCAGCTGGAGAAATGTGGAACAAGTCAACACAGACACAGAAACCGCC 432  
DB 123 LeuAlaGlnTyrGlyThrValGluAsnValGluGlnValAsnThrAspThrGluThrAla 142  
QY 433 GTTGTCAAGCTCACATATGCACACAGACAGAGACAAAATAGCCATGGAGAGCTAAGC 492  
DB 143 ValValAsnValThrTyrAlaThrArgGluGluAlaLysIleAlaMetGluLysLeuSer 162  
QY 493 GGGCATCAGTTTGAACAATCTCTTCAAGATTTCTACATCTCCCGATCGGAGGTGAGC 552  
DB 163 GlyHisGlnPheGluAsnTyrSerPheLysIleSerTyrIleProAspGluGluValSer 182  
QY 553 TCCCTCTGCCCTCAGCGAGCGCGCGAGGTGGGGACCACTCTTCCCGGAGCAAGGCCAC 612

DB 183 SerProSerProGlnArgAlaGlnArgGlyAspHisSerSerArgGluGlnGlyHis 202  
QY 613 GCCCTTGGGGGCACTTCTCAGGCCAGACAGATTGATTTCCTCCGCTGCGGATCTGGTCCCC 672  
DB 203 AlaProGlyGlyThrSerGlnAlaArgGlnIleAspPheProLeuArgIleLeuValPro 222  
QY 673 ACCCAGTTTGTGTGGCCATCATCGAAAGAGGAGGCTTGACCATAAAGAACATCACTAAG 732  
DB 223 ThrGlnPheValGlyAlaIleIleGlyLysGluGlyLeuThrIleLysAsnIleThrLys 242  
QY 733 CAGACCCAGTCCCGGTAGATATCATAGAAAAGAGAAGTCTGGAGCTGCAGAGAGCCT 792  
DB 243 GlnThrGlnSerArgValAspIleHisArgLysGluAsnSerGlyAlaAlaGluLysPro 262  
QY 793 GTCCACCATCATGCCACCCAGAGGGAGCTTCTAGCATGCCGATGATTTCTGAAATC 852  
DB 263 ValThrIleHisAlaThrProGluGlyThrSerGluAlaCysArgMetIleLeuGluIle 282  
QY 853 ATGCAGAAAGAGGAGGAGATGAGCAAACTAGCCGAAGAGATTCCTCTGAAAATCTTGCA 912  
DB 283 MetGlnLysGluAlaAspGluThrLysLeuAlaGluGluIleProLeuLysIleLeuAla 302  
QY 913 CACAATGGCTTGTGTGAAGACTGATTGGAAGAGGACGAGAAATTTGAGAAATTTGAA 972  
DB 303 HisAsnGlyLeuValGlyArgLeuIleGlyLysGluGlyArgAsnLeuLysLysIleGlu 322  
QY 973 CATGAAACAGGAGGACCAAGATAAACAATCTCATCTTTGAGGATTTGAGCATATACACCG 1032  
DB 323 HisGluThrGlyThrLysIleThrIleSerSerLeuGlnAspLeuSerIleTyrAsnPro 342  
QY 1033 GAAAGAACCATCACTGTGAAGGCGACAGTTGAGGCTGTGCGAGTCTGAGATAGAGATT 1092  
DB 343 GluArgThrIleThrValLysGlyThrValGluAlaCysAlaSerAlaGluIleGluIle 362  
QY 1093 ATGAAGAAGCTCGTGAAGCCTTTGAAATGATGCTGGTGTGTTAAACAACAAGCCCAAT 1152  
DB 363 MetLysLysLeuArgGluAlaPheGluAsnAspMetLeuAlaValAsnGlnAlaAsn 382  
QY 1153 CTGATCCAGAGGTGAACTCAGCGCACTTGGCATCTTTTCAACAGGAGCTGCGGTGCTA 1212  
DB 383 LeuIleProGlyLeuAsnLeuSerAlaLeuGlyIlePheSerThrGlyLeuSerValLeu 402  
QY 1213 TCTCACAGCAGGCGCCCGCGGAGCTCCCGCGCTGCCCTTACACCCCTTCACTACC 1272  
DB 403 SerProProAlaGlyProArgGlyAlaProProAlaAlaProTyrHisProPheThrThr 422  
QY 1273 CACTCCGGATCTTCTCAGCCTGTACCCCATCACAGTTTGGCCCTTCCCGCATCAT 1332  
DB 423 HisSerGlyTyrPheSerSerLeuTyrProHisGlnPheGlyProPheProHisHis 442  
QY 1333 CACTCTTATCCAGAGCAGGAGATTGTGAATCTTTCATCCCAACCCAGCGCTGGGGCC 1392  
DB 443 HisSerTyrProGluGlnGluIleValAsnLeuPheIleProThrGlnAlaValGlyAla 462  
QY 1393 ATCATCGGAAAGAGGGGCGCACACATCAACAGCTGGCGAGATTTCGCGGAGCCCTCTATC 1452  
DB 463 IleIleGlyLysLysGlyAlaHisIleLysGlnLeuAlaArgPheAlaGlyAlaSerIle 482  
QY 1453 AAGATTGCCCTCGGAAAGGCCGAGACGTACAGCAAGAGATGGTTCATCATCACCGGCGCA 1512  
DB 483 LysIleAlaProAlaGluGlyProAspValSerGluArgMetValIleIleThrGlyPro 502  
QY 1513 CCGGAAGCCCATGTTCAAGGCCAGGAGCGGATCTTTGGGAACTTGAAGAGAGAAATTC 1572  
DB 503 ProGluAlaGlnPheLysAlaGlnGlyArgIlePheGlyLysLeuLysGluGluAsnPhe 522  
QY 1573 TTTAAACCCCAAGAGAGTGAAGCTGGAAGCGCATATCAGAGTGCCTCTTTCACAGCT 1632  
DB 523 PheAsnProLysGluGluValLysLeuGlu\*\*\*HisIleArgValProSerSerThrAla 542  
QY 1633 GGCGGGTGTATGGCAAGGTGGCAAGCCCTGAACGAACTGCAGAACTTAACAGTCA 1692

Db 543 GlyArgValIleGlyLysGlyLysThrValAsnGluLeuGlnAsnLeuThrSerAla 562

QY 1693 GAAGTCATCTGCTCGTGCCTGACCCAAACGCCAGATGAAATCAGGAATGATCGTCAGAATT 1752

Db 563 GluValIleValProArgAspGlnThrProAspGluAsnGluGluValIleValArgIle 582

QY 1753 ATCGGGCACTCTTCTGCTAGCCAGACTGCACAGCGGAGATCAGGGAATGTCACACAG 1812

Db 583 IleGlyHisPheIleAlaSerGlnThrAlaGlnArgLysIleArgGluIleValGlnGln 602

QY 1813 GTGAAGCAGCAGGAGCAGAGAATACCTCTCAGGAGTCGCCCTCACAGCGCAGCAAG 1866

Db 603 VallysGlnGlnGluGlnLysTyrProGlnGlyValAlaSerGlnArgSerLys 620

RESULT 4

ABG06795

ID ABG06795 standard; protein; 594 AA.

XX ABG06795;

XX AC

XX DE

DT 13-FEB-2002 (first entry)

XX Novel human diagnostic protein #6786.

DE Human; chromosome mapping; gene mapping; gene therapy; forensic;

XX food supplement; medical imaging; diagnostic; genetic disorder.

XX Homo sapiens.

XX WO200175067-A2.

XX 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US008631.

XX 31-MAR-2000; 2000US-00540217.

PR 23-AUG-2000; 2000US-00649167.

XX (HYSE-) HYSEQ INC.

PA Drmanac RT, Liu C, Tang YT;

PI WPI; 2001-639362/73.

DR N-PSDB; AAS70982.

XX New isolated polynucleotide and encoded polypeptides, useful in

PT diagnostics, forensics, gene mapping, identification of mutations

PT responsible for genetic disorders or other traits and to assess

PT biodiversity.

XX Claim 20; SEQ ID NO 37154; 103pp; English.

PS The invention relates to isolated polynucleotide (I) and polypeptide (II)

CC sequences. (I) is useful as hybridisation probes, polymerase chain

CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,

CC and in recombinant production of (II). The polynucleotides are also used

CC in diagnostics as expressed sequence tags for identifying expressed

CC genes. (I) is useful in gene therapy techniques to restore normal

CC activity of (II) or to treat disease states involving (II). (II) is

CC useful for generating antibodies against it, detecting or quantitating a

CC polypeptide in tissue, as molecular weight markers and as a food

CC supplement. (II) and its binding partners are useful in medical imaging

CC of sites expressing (II). (I) and (II) are useful for treating disorders

CC involving aberrant protein expression or biological activity. The

CC polypeptide and polynucleotide sequences have applications in

CC diagnostics, forensics, gene mapping, identification of mutations

CC responsible for genetic disorders or other traits to assess biodiversity

CC and to produce other types of data and products dependent on DNA and

CC amino acid sequences. ABG0010-ABG0377 represent novel human diagnostic

CC amino acid sequences of the invention. Note: The sequence data for this

CC patent did not appear in the printed specification, but was obtained in

CC electronic format directly from WIPO at

CC ftp.wipo.int/pub/published\_pct\_sequences

XX SQ Sequence 594 AA;

Alignment Scores:

Pred. No.: 5,51e-255 Length: 594

Score: 2917.50 Matches: 577

Percent Similarity: 93.08% Conservative: 1

Best Local Similarity: 92.91% Mismatches: 0

Query Match: 47.66% Indels: 43

DB: 4 Gaps: 1

US-09-270-437D-6 (1-3412) x ABG06795 (1-594)

QY 4 AGCGAGGAGGCGCAGAGCGCCGGTACCGGGCGGGGAGCGCGGGCTCTCGGGGAAG 63

Db 17 AsnGlyGlyGlyGluGluArgArgValProGlyArgGlySerArgGlyLeuSerGlyLys 36

QY 64 AGACGGATGATGAACAAGCTTTACATCGGGAACCTGAGCCCGCCCTCACCGCGGAC 123

Db 37 ArgGwMetMetAsnLysLeuTyrlleGlyAsnLeuSerProAlaValThrAlaAsp 56

QY 124 CTCGGCAGCTCTTTGGGGACAGGAAGCTGCCCTCGCGGACAGGTCTCTGCTGAAGTCC 183

Db 57 LeuArgGlnLeuPheGlyAspArgLysLeuProLeuAlaGlyGlnValLeuLeuLysSer 76

QY 184 GGCTAGCCTTCTGTGACTACCCCGACACAGAACTGGGGCCATCCGCGCATCGAGACCTC 243

Db 77 GlyTyrAlaPheValAspTyrProAspGlnAsnTrpAlaIleArgAlaIleGluThrLeu 96

QY 244 TCGGGTAAAGTGAATTGCATGGGAAATCATCGAAGTTGATTACTCAGTCTCTAAAG 303

Db 97 SerGlyLysValGluLeuHisGlyLysIleMetGluValAspTyrSerValSerLysLys 116

QY 304 CTAAGGAGCAGGAAAATTCAGATTTCGAAACATCCCTCCTCACCTGAGTGGAGTGTG 363

Db 117 LeuArgSerArgLysIleGlnIleArgAsnIleProProHisLeuGlnTrpGluValLeu 136

QY 364 GATGGACTTTTGGCTCAATATGGGACAGTGGAGAATGTGGACAAGTCAACACAGACACA 423

Db 137 AspGlyLeuLeuAlaGlnTyrGlyThrValGluAsnValGluGlnValAsnThrAspThr 156

QY 424 GAAACCGCGCTTGTCAACGTTCACATATGCAACAGAGAGAGAAACAAATAGCATCGAG 483

Db 157 GluThrAlaValValAsnValThrTyrAlaThrArgGluGluAlaLysIleAlaMetGlu 176

QY 484 AAGCTAAGCGGCGCATCAGTTTGAGAACTACTCTCAAGATTTCTACATCCCGGATGAA 543

Db 177 LysLeuSerGlyHisGlnPheGluAsnTyrSerPheLysIleSerTyrlleProAspGlu 196

QY 544 GAGTGAGTCTCCCTTCGCCCCCTCAGCGAGCCCGAGCGGGGACCACTCTTCCCGGGAG 603

Db 197 GluValSerSerProSerProGlnArgAlaGlnArgGlyAspHisSerSerArgGlu 216

QY 604 CAAAGGCCAGCCCTCGGGGSCACTTCTCAGGCCAGACAGATTGATTTCGCGCTCGCGATC 663

Db 217 GlnGlyHisAlaProGlyGlyThrSerGlnAlaArgGlnIleAspPheProLeuArgIle 236

QY 664 CTGGTCCCCACCCAGTTTGTGTGTCATCATCGAAAGAGGGCTTGACCATAAAGAAC 723

Db 237 LeuValProThrGlnPheValGlyAlaIleGlyLysGluGlyLeuThrIleLysAsn 256

QY 724 ATCACTAAGCAGACCCAGTCCCGGGTAGATATCCATAGAAAGAGAACCTCTGAGCTGCA 783

Db 257 IleThrLysGlnThrGlnSerArgValAspIleHisArgLysGluAsnSerGlyAlaAla 276

QY 784 GAGAAGCTGTCAACCATCCATGCCACCCAGAGGGGACTTCTGAAGCATGCCCATGATT 843

Db 277 GluLysProValThrIleHisAlaThrProGluGlyThrSerGluAlaCysArgMetIle 296

QY 844 CTTTAAATCATGCAGAAAGAGGAGATGAGACCAAACTAGCCGGAAGAGATTTCCTCTGAA 903

Db 297 LeuGluIleMetGlnLysGluAlaAspGluThrLysLeuAlaGluGluLeuProLeuLys 316

QY 904 ATCTTGGCACCAATGGCTTGGTGAAGACTGATTGAAAGAGGAGGAGCAAAATTTGAAG 963  
Db |||||||  
QY 317 IleuAlaHisAsnGlyLeuValGlyArgLeuIleGlyLysGluGlyArgAsnLeuLys 336  
Db |||||||  
QY 964 AAAATTGAACATGAACAGGAGCAAGATAACAATCTCATCTTTGACGATTGTGACATA 1023  
Db |||||||  
QY 337 LysIleGluHisGluThrGlyThrLysIleThrLysSerSerLeuGlnAspLeuSerIle 356  
QY 1024 TACACCCGGAAGAACCATCACTGTGAAGGACAGTGTAGGCTGTGCCAGTCTCTAG 1083  
Db |||||||  
QY 357 TyrAsnProGluArgThrIleThrValLysGlyThrValGluAlaCysAlaSerAlaGlu 376  
QY 1084 ATAGAGATTATGAAGAAGCTGGTGAAGCCTTTGAAATGATATGCTGGCTGTAAACAA 1143  
Db |||||||  
QY 377 IleGluIleMetLysLysLeuArgGluAlaPheGluAsnAspMetLeuAlaValAsn--- 395  
QY 1144 CAAGCCAATCTATCCCAAGGTTGAACCTCAGCGCACTTGGCATCTTTTCAACAGGACTG 1203  
Db |||||||  
QY 395 ----- 395  
QY 1204 TCCGTGCTATCTCCACAGCAGGCGCCGCGAGCTCCCCCGCTGCCCTTACCACCCC 1263  
Db |||||||  
QY 395 ----- 395  
QY 1264 TTCACCTACCCACTCCGATATCTTCCAGCCTGTACCCCCATCACCAGTTTGGCCCGTTC 1323  
Db |||||||  
QY 396 -----ThrHisSerGlyTyrPheSerSerLeuTyrProHisHisGlnPheGlyProPhe 413  
QY 1324 CCGATCATCATCTTATCCAGAGCAGGAGATTGTGAATCTTTCATCCCAACCCAGGCT 1383  
Db |||||||  
QY 414 ProHisHisSerTyrProGluGlnGluIleValAsnLeuPheIleProThrGlnAla 433  
QY 1384 GTGGCGCCATCATCGGAAGAGGGGCACACATCAACAGCTGCGCAGATTGCCCGGA 1443  
Db |||||||  
QY 434 ValGlyAlaIleIleGlyLysGlyAlaHisIleLysGlnLeuAlaArgPheAlaGly 453  
QY 1444 GCCTCTATCAAGATTGCCCTCGGAAGCCGACAGTCAGCAAGAGATGTCATCATC 1503  
Db |||||||  
QY 454 AlaSerIleLysIleAlaProAlaGluGlyProAspValSerGluArgMetValIleIle 473  
QY 1504 ACCGGGCCACCGGAAGCCAGTTCAAGCCGACGAGCGATCTTTGGGAATCTGGAAGAG 1563  
Db |||||||  
QY 474 ThrGlyProProGluAlaGlnPheLysAlaGlnGlyArgIlePheGlyLysLeuLysGlu 493  
QY 1564 CAAAACTCTTTAAACCCCAAGAAAGTGAAGCTGGAAGCGCATATCAGAGTGCCCTCT 1623  
Db |||||||  
QY 494 GluAsnPhePheAsnProLysGluGluValLysLeuGluAlaHisIleArgValProSer 513  
QY 1624 TCCACAGCTGCCCGGTGATTGGCAAGGTGGCAAGCCGTGAACGAACCTGCAGAACTTA 1683  
Db |||||||  
QY 514 SerThrAlaGlyArgValIleGlyLysGlyLysThrValAsnGluLeuGlnAsnLeu 533  
QY 1684 ACCAGTCAGAGTCATCGTCTGTCACCAACCCAGATGAAATGAGGAAGTATC 1743  
Db |||||||  
QY 534 ThrSerAlaGluValIleValProArgAspGlnThrProAspGluAsnGluGluValIle 553  
QY 1744 GTCAGAATTATCGGCACTTCTTCTAGCCAGCTGCACAGCGCAAGATCAGGGAATTT 1803  
Db |||||||  
QY 554 ValArgIleIleGlyHisPhePheAlaSerGlnThrAlaGlnArgLysIleArgGluIle 573  
QY 1804 GTACACAGGTGAAGCAGCAGAGCAGAAATACCTCTAGGAGTGCCTCAGCGCAGC 1863  
Db |||||||  
QY 574 ValGlnGlnValLysGlnGlnGlnLysTyrProGlnGlyValAlaSerGlnArgSer 593  
QY 1864 AAG 1866  
Db |||||  
QY 594 Lys 594  
RESULT 5  
ABG96346  
ID ABG96346 standard; protein; 556 AA.  
XX  
AC ABG96346;

XX 11-DEC-2002 (first entry)  
DT Human ovarian cancer marker M452.  
XX  
DE Human; ovarian cancer; marker; cancer; familial history; brain disorder;  
KW central nervous system disorder; bacterial meningitis; viral meningitis;  
KW Alzheimer's disease; Parkinson's disease; cerebral edema; hydrocephalus;  
KW brain herniation; inflammation; encephalitis; testicular disorder;  
KW nontuberculous granulomatous orchitis; connective tissue disorder;  
KW heart disorder; ischaemic heart disease; atherosclerosis; neoplasm;  
KW histological type; carcinogenic; ovarian cancer marker.  
XX Homo sapiens.  
OS  
XX WO200271928-A2.  
PN 19-SEP-2002.  
XX  
PD 14-MAR-2002; 2002WO-US007826.  
PF  
XX 14-MAR-2001; 2001US-0276025P.  
PR 14-MAR-2001; 2001US-0276026P.  
PR 10-AUG-2001; 2001US-0311732P.  
PR 19-SEP-2001; 2001US-0323580P.  
PR 26-SEP-2001; 2001US-0324967P.  
PR 26-SEP-2001; 2001US-0325102P.  
PR 26-SEP-2001; 2001US-0325149P.  
XX (MILL-) MILLENNIUM PHARM INC.  
PA  
PI Monahan JE, Gannavarapu M, Hoersch S, Kamatkar S, Kovatis SG;  
PI Meyers RE, Morrissey MP, Olandt PJ, Sen A, Vieby PO, Mills GB;  
PI Basc RC, Lu K, Schmandt RE, Zhao X, Glatt K;  
XX WPI; 2002-723277/78.  
DR N-PSDB; ABS76442.  
XX  
PT Assessing whether a patient is afflicted with ovarian cancer, useful in  
PT assessing the stage or progression of the disease, comprises comparing  
PT the expression level of a cancer marker in a sample from a patient and  
PT from a non cancer patient.  
PS Disclosure; Page 263-264; 481pp; English.  
XX The present invention relates to a new method for assessing whether a  
CC patient is afflicted with ovarian cancer. The method involves comparing  
CC the expression level of a marker in a patient sample and the normal level  
CC of expression of the marker in a control non-ovarian cancer sample, where  
CC the marker is selected from 363 cancer markers described in the  
CC specification. The method of the invention is useful in diagnosing or  
CC characterising cancer, in detecting the presence of cancer as early as  
CC possible, and the recurrence of ovarian cancer. The method may also be of  
CC particular use with patients having an enhanced risk of developing  
CC ovarian cancer (e.g. patients having a familial history of ovarian  
CC cancer). The cancer markers may be used in the management and treatment  
CC of e.g. brain and central nervous system disorders (e.g. bacterial and  
CC viral meningitis, Alzheimer's disease or Parkinson's disease), brain  
CC disorders (e.g. cerebral oedema, hydrocephalus or brain herniations),  
CC inflammations (e.g. bacterial or viral meningitis or encephalitis),  
CC testicular disorders (e.g. nontuberculous granulomatous orchitis),  
CC connective tissue disorders, or heart disorders (e.g. ischaemic heart  
CC disease or atherosclerosis). The compositions and methods may also be  
CC used in assessing the histological type of neoplasm associated with  
CC ovarian cancer, monitoring the progression of ovarian cancer, determining  
CC whether ovarian cancer has metastasized or is likely to metastasize,  
CC selecting a composition for inhibiting ovarian cancer, assessing the  
CC ovarian carcinogenic potential of a compound, or inhibiting ovarian  
CC cancer or at risk of developing ovarian cancer. The present amino acid  
CC sequence represents one of the ovarian cancer markers described in the  
XX invention  
XX Sequence 556 AA;  
SQ



## Alignment Scores:

Pred. No.: 6,19e-245 Length: 556  
Score: 2806.50 Matches: 556  
Percent Similarity: 92.82% Conservative: 0  
Best Local Similarity: 92.82% Mismatches: 0  
Query Match: 45.85% Indels: 43  
DB: 5 Gaps: 1

US-09-270-437D-6 (1-3412) x ABG96346 (1-556)

```
QY 70 ATGATGAACAGCTTACATCGGGAACCTGAGCCCGCGCTCAGCCGCGAGCCTCCGG 129
Db 1 MetMetAsnLysLeuYrilleGlyAsnLeuSerProAlaValThrAlaAspLeuArg 20
QY 130 CAGCTCTTTGGGACAGGAAGCTGCCCTCGCGGGCAGGTCCTGTGTAAGTCGGGTAC 189
Db 21 GlnLeuPheGlyAspArgLysLeuProLeuAlaGlyGlnValLeuLysSerGlyTyr 40
QY 190 GCCTTTCGTGACTACCCCGACCAAGACTGGGCCATCCGCGCCATCGAGCCCTCTCGGT 249
Db 41 AlaPheValAspTyrProAspGlnAsnTrpAlaIleArgAlaIleGluThrLeuSerGly 60
QY 250 AAGTGTGAATTGCATGGGAATAATCATGCAAGTTGATTACTCAGTCTCTAAAAAGCTAAGG 309
Db 61 LysValGlnLeuHisGlyLysIleMetGluValAspTyrSerValSerLysLysLeuArg 80
QY 310 AGCAGGAAATTCAGATTCGAAACATCCCTCCTCACCTGAGTGGGAGGTGTGGATGGA 369
Db 81 SerArgLysIleGlnIleArgAsnIleProProHisLeuGlnTrpGluValLeuAspGly 100
QY 370 CTTTTGGCTCAATATGGGACAGTGGAGAATGTGGAACAAGTCAACACAGACAGAAACC 429
Db 101 LeuLeuAlaGlnTyrGlyThrValGluAsnValGluGlnValAsnThrAspThrGluThr 120
QY 430 GCCCTTCTCAAGTCACATATGCAACAGAGAGAGCAAAAAATAGCCATGGAGAAGCTA 489
Db 121 AlaValValAsnValThrTyrAlaThrArgGluGluAlaLysIleAlaMetGluLysLeu 140
QY 490 AGCGGGCATCAGTTTGAGAACTACTCCTTCAAGATTTCTTACATCCCGGATGAGAGGTG 549
Db 141 SerGlyHisGlnPheGluAsnTyrSerPheLysIleSerTyrIleProAspGluGluVal 160
QY 550 AGCTCCCTTCGCCCTCAGCAGGCCAGCGTGGGACCACTCTTCCCGGGAGCAGGC 609
Db 161 SerSerProSerProProGlnArgAlaGlnArgGlyAspHisSerSerArgGluGlnGly 180
QY 610 CAGCCCTCGGGGCACTTCTCAGGCCAGACAGATTGATTTCCCGCTCGGATCCTGGTC 669
Db 181 HisAlaProGlyGlyThrSerGlnAlaArgGlnIleAspPheProLeuArgIleLeuVal 200
QY 670 CCACCCAGTTTGTGTGGCATCATCGGAAGAGGGCTTGACCATPAAAGAACATCACT 729
Db 201 ProThrGlnPheValGlyAlaIleIleGlyLysGluGlyLeuThrIleLysAsnIleThr 220
QY 730 AAGCAGACCCAGTCCCGGTAGATATCCATAGAAAGAGACTCTGAGCTGCAGAGAAG 789
Db 221 LysGlnThrGlnSerArgValAspIleHisArgLysGluAsnSerGlyAlaAlaGluLys 240
QY 790 CTTGTACACATCCATGCCACCCAGAGGGGACTTCTGAAGCATGCCGATGATTCTTGAA 849
Db 241 ProValThrIleHisAlaThrProGluGlyThrSerGluAlaCysArgMetIleLeuGlu 260
QY 850 ATCATGCAAGAAGGAGATGAGACCAAACTAGCCGAAGAGATTCTCTGAAATCTTG 909
Db 261 IleMetGlnLysGluAlaAspGluThrLysLeuAlaGluGluIleProLeuLysIleLeu 280
QY 910 GCACACAATGCTTGGTGGAGACTGATTGGAAAAGAGCGCAGAAATTTGAAGAAATT 969
Db 281 AlaHisAsnGlyLeuValGlyArgLeuIleGlyLysGluGlyArgAsnLeuLysLysIle 300
QY 970 GAACATGAACAGGGACCAAGATAACAATCTCATCTTTTGAGGATTTGAGCATATCAAC 1029
```

```
Db 301 GluHisGluThrGlyThrLysIleThrIleSerSerLeuGlnAspLeuSerIleTyrAsn 320
QY 1030 CCGAAAGAACCATCATCTGTGAAGGGCACAGTTGTAGCCCTGTGCCAGTGTGTGAGATAGAG 1089
Db 321 ProGluArgThrIleThrValLysGlyThrValGluAlaCysAlaSerAlaGluIleGlu 340
QY 1090 ATTATGAAGAAGCTGCGTGAAGGCTTTGAAAAATGATATGTGCTGTATTACCAACAAGCC 1149
Db 341 IleMetLysLysLeuArgGluAlaPheGluAsnAspMetLeuAlaValAsn----- 357
QY 1150 AATCTGATCCACAGGGTTGAACCTCAGCGCACTTGGCATCTTTTCAACAGGACTGTCCGTG 1209
Db 357 ----- 357
QY 1210 CTATCTCCACAGAGGGGCGCGGAGCTCCCCCGCTGCCCCCTACCAACCCCTTCACT 1269
Db 357 ----- 357
QY 1270 ACCACTCCGGATACTTCTCCAGCTGTACCCCTATCACCAGTTTGGCCCGTTCGCGCAT 1329
Db 358 ThrHisSerGlyTyrPheSerSerLeuTyrProHisGlnPheGlyProPheProHis 377
QY 1330 CATCACTCTTATCCAGACAGGAGATTGTAATCTCTTATCCCAACCCAGGCTGTGGGC 1389
Db 378 HisHisSerTyrProGluGlnGluIleValAsnLeuPheIleProThrGlnAlaValGly 397
QY 1390 GCATCATCGGGAAGGGGGCACACATCAACAGCTGGCAGATTGCCGAGGCTCT 1449
Db 398 AlaIleIleGlyLysGlyAlaHisIleLysGlnLeuAlaArgPheAlaGlyAlaSer 417
QY 1450 ATCAAGATTCCTCCCTCGGAGGCCAGCGTCAGCGAAGGATGTCATCATCAGCGG 1509
Db 418 IleLysIleAlaProAlaGluGlyProAspValSerGluArgMetValIleIleThrGly 437
QY 1510 CCACCGGAAGCCAGTTCAAGGCCAGGGACGATCTTTGGGAAACTGAAAGAGGAAAC 1569
Db 438 ProProGluAlaGlnPheLysAlaGlnGlyArgIlePheGlyLysLeuLysGluGluAsn 457
QY 1570 TCTTTTAAACCCAAAGAAAGTGAAGCTGGAAGCGCATATCAGAGTGCCTCTTCCACA 1629
Db 458 PhePheAsnProLysGluGluValLysLeuGluAlaHisIleArgValProSerSerThr 477
QY 1630 GCTGGCGGGTGAATTGGCAAGGTGGCAAGCCGTGAACCAACTGCAGAGACTTAACCACT 1689
Db 478 AlaGlyArgValIleGlyLysGlyLysThrValAsnGluLeuGlnLeuThrSer 497
QY 1690 GCAGAACTCATCTGCTCGTCCGACCAACCCAGATGAAATGAGGAAGTGCATCGTCAGA 1749
Db 498 AlaGluValIleValProArgAspGlnThrProAspGluAsnGluGluValIleValArg 517
QY 1750 ATTATCGGGCACTTCTTTTGTAGCCAGACTGCACAGCGCAAGATCAGGGAAATTTGACAA 1809
Db 518 IleIleGlyHisPhePheAlaSerGlnThrAlaGlnArgLysIleArgGluIleValGln 537
QY 1810 CAGGTGACGACGAGGACAGCAATACCTCAGGGAGTCCCTCAGACGACGACGACG 1866
Db 538 GlnValLysGlnGluGlnLysTyrProGlnGlyValAlaSerGlnArgSerLys 556
```

## RESULT 6

ABU89799  
ID ABU89799 standard; protein; 555 AA.

AC ABU89799;

XX 10-JUL-2003 (first entry)

XX Novel human protein NOV14a.

XX Human; cytostatic; DAPK3-Agonist; DAPK3-Antagonist; cancer; NOV.

XX Homo sapiens.

XX W02003031571-A2.

XX PD 17-APR-2003.  
 XX PF 02-OCT-2002; 2002WO-US031357.  
 XX PR 05-OCT-2001; 2001US-0327454P.  
 PR 09-OCT-2001; 2001US-0327917P.  
 PR 09-OCT-2001; 2001US-0328029P.  
 PR 09-OCT-2001; 2001US-0328056P.  
 PR 12-OCT-2001; 2001US-0328849P.  
 PR 15-OCT-2001; 2001US-0329414P.  
 PR 17-OCT-2001; 2001US-0330142P.  
 PR 22-OCT-2001; 2001US-0341058P.  
 PR 24-OCT-2001; 2001US-0343629P.  
 PR 29-OCT-2001; 2001US-0349575P.  
 PR 01-NOV-2001; 2001US-0346357P.  
 PR 25-JUN-2002; 2002US-0391342P.  
 PR 01-OCT-2002; 2002US-0026244S.  
 XX PR (CURA-) CURAGEN CORP.  
 XX PI Alsbrook JP, Burgess CE, Catterton E, Chant JS, Chaudhuri A;  
 PI Edinger SR, Gerlach VL, Giot L, Gorman L, Guo X, Kekuda R;  
 PI Mezes PS, Millet I, Cui CE, Patturajan M, Rieger DK, Spytek KA;  
 PI Taupier RJ, Zernusen BD, Zhong H, Zhong M;  
 XX WPI; 2003-381704/36.  
 DR N-PSDB; ACA90176.  
 XX PT New DAPK3 polypeptide, useful for preparing a composition for treating or  
 PT preventing e.g., cancer.  
 XX PS Claim 2; Page 129; 253pp; English.  
 XX CC The invention describes an isolated polypeptide comprising any of 33 90-  
 CC 1273 amino acid sequences (I) given in the specification or its mature  
 CC form, a sequence that is at least 95 % identical to (I), or a sequence  
 CC comprising one or more conservative substitutions in the amino acid  
 CC sequence of (I). The polypeptide is useful for preparing a composition  
 CC for treating or preventing e.g. cancer. This is the amino acid sequence  
 CC of a novel human NOV protein  
 XX SQ Sequence 555 AA;  
 Alignment Scores:  
 Pred. No.: 6,01e-222 Length: 555  
 Score: 2553,00 Matches: 507  
 Percent Similarity: 87.81% Conservative: 19  
 Best Local Similarity: 84.64% Mismatches: 29  
 Query Match: 41.71% Indels: 44  
 DB: 6 Gaps: 2  
 US-09-270-437D-6 (1-3412) x ABU89799 (1-555)  
 QY 70 ATGATGAACAGCTTTACATCGGAACTGAGCCCGCCGCTCAGCCGACGACCTCGG 129  
 DB 1 MetMetAsnLysLeuPheIleGlyAsnLeuSerProAlaValThraAlaGluAspLeuArg 20  
 QY 130 CAGCTCTTTGGGACAGGAAGCTCCCTCGGGGACAGGTCCTGCTGAAGTCCGGCTAC 189  
 DB 21 GlnLeuPheGlyAspArgLysLeuProLeuAlaGlyGlnValLeuLysSerArgTyr 40  
 QY 190 GCCTTCGTGGACTACCCGACAGAACTGGGCCCATCCGCGCATCGAGACCTCTCGGGT 249  
 DB 41 AlaPheValAspTyrProAspGlnAsnTrpAlaIleArgThrIleGluThrLeuSerGly 60  
 QY 250 AAAGTGAATGCATGGGAAATCATGGAAGTTGATTTACTCAGTCTCTAAAAGCTAAGG 309  
 DB 61 GlnValGluLeuHisGlyLysIleMetGluValAspTyrSerValSerIleLysLeuArg 80  
 QY 310 AGCAGGAAATTCAGATTCGAAACATCCCTCCTCAGCTGAGTGGGAGGTGTGGATGGA 369  
 DB 81 SerArgAsnIleProIleArgAsnIleProProHisLeuGlnTrpGluValLeuAspGly 100

QY 370 CTTTGGCTCAATATGGGACAGCTGAGAAATCTGGAACAAGTCAACACACACACACAAACC 429  
 DB 101 LeuLeuAlaGlnTyrGlyThrValGluAsnValGluGlnValAsnThrAspThrGluThr 120  
 QY 430 GCCGTTGTCAACGCTCACATATGCAACAGAGAAGCAAAATAGCCATGAGAGAGCTA 489  
 DB 121 AlaValValAsnValThrTyrAlaThrLysGluGluValLysIleAlaMetLysLysLeu 140  
 QY 490 AGCGGCGATCAGTTTGAGAACTATCTCTTCAAGATTCTCTACATCCCGATGAAGGTG 549  
 DB 141 SerGlyHisGlnPheGluAsnHisTyrPheLysIleSerTyrIleProAspAspGluVal 160  
 QY 550 AGCTCCCTCCGCGCCCTCAGCGCCAGCGTGGGAGCACTCTTCCCGGAGCAAGGC 609  
 DB 161 SerCysProSerProGlnArgAlaGlnArgGlyAspHisSerSerTyrGluGlnGly 180  
 QY 610 CACGCCCTGGGGCAGCTTCTCAGGCCAGACAGATTGATTTCCCGCTCGGATCCTGTC 669  
 DB 181 GlnAlaProGlyGlySerSerGlnAlaArgGlnIleAspPheProLeuArgValLeuPhe 200  
 QY 670 CCCACCCAGTTTGTGTGGTCCATCATCGAAAGAGGGCTTGACCATTAAGACATCACT 729  
 DB 201 ProThrGlnPheValGlyAlaIleGlyLysGluGlyLeuThrIleLysAsnIleThr 220  
 QY 730 AAGCAGACCCAGTCCCGGTAGATATCCATAGAAAGAGAACTCTGGAGCTGCAGAGAG 789  
 DB 221 LysGlnSerArgSerArgValAspIleTyrArgGlnGlnAsnSerArgAlaAlaGluLys 240  
 QY 790 CTTGTACCATTCATCCACCCAGAGGGGAGCTTCTGAAGCATGCCGATGATCTTGA 849  
 DB 241 ProValThrMetHisAlaThrProGluGlyThrSerGluAlaCysArgMetIleLeuGlu 260  
 QY 850 ATCATGCAAGAGAGGAGATGACACAACTAGCCGAAAGAGATTCCTCTGAAATCTTG 909  
 DB 261 IleMetGlnLysGluAlaAspGluAlaLysLeuAlaGluGluIleProLeuLysIleLeu 280  
 QY 910 GCACCAATGCTTGTGTGAAGACTGATTCGAAAGAGAGCAAAATTTCAAGAAATTT 969  
 DB 281 AlaHisGlnGlyLeuValGlyArgLeuIleGlyLysGluGlyArgAsnLeuLysAsn 300  
 QY 970 GAACATGAACACAGGACCAAGATAACATCTCTTTTGCAGGATTTGAGCATATACAA 1029  
 DB 301 GluHisGluThrGlyThrLysIleThrIleSerSerGlnAspLeuSerIleTyrAsn 320  
 QY 1030 CCGAAAGAACCATCACTGTCAAGGGACAGTTGAGGCTGTGCGAGTGTGAGATAGAG 1089  
 DB 321 ProGluArgThrIleThrValLysGlyThrValGluValCysAlaSerAlaGluIleGlu 340  
 QY 1090 ATTATGAAGAGCTCGGTGAGGCTTTGAAATGATATGCTGGCTGTTAAACCAACAGCC 1149  
 DB 341 IleMetLysLysLeuArgGluAlaPheGluAsnAspThrLeuThrValAsn----- 357  
 QY 1150 AATCTGATCCAGGGTTGAACCTCAGCGCACTTGGCATCTTTTCAACAGGAGCTGTCCGTG 1209  
 DB 357 ----- 357  
 QY 1210 CTATCTCCACAGAGGGCCCGGGAGCTCCCGCGCTCCCGCTTACCACCCCTTCACT 1269  
 DB 357 ----- 357  
 QY 1270 ACCCACTCCGATATCTTCTCCAGCTGTACCCCATCACAGTTTGGCCGCTTCCCGCAT 1329  
 DB 358 ThrHisPheGlyTyrPheSerSerLeuTyrProHisArgGlnPheGlyProPheProHis 377  
 QY 1330 CATCACTCTTATCCAGAGCAGGAGATGTGATCTCTTCACTCCCAACCCAGGCTGTGGC 1389  
 DB 378 HisHisSerTyrProGluGlnGluIleValAsnLeuPheIleProThrGlnGlyValGly 397  
 QY 1390 GCCATCATCGGAAAGAGGGGGCACATCAACAGCTGGGAGATTCGCGGAGCTCT 1449  
 DB 398 AlaIleIleGlyLysLysGlyAlaHisIleLysGlnLeuAlaArgPheValGlyAlaSer 417

QY 1450 ATCAAGATTCCCTCGGAGGCCACAGCTCAGCGAAAGGATGGTTCATCATCACCGGG 1509  
 Db ||||||| ||||||| ||| : : : ||| ||||||| ||||||| |||||||  
 QY 418 IlelyslleAlaProAlaArgSerPro--LeuArgGlnArgLysValIleIleThrTrp 436  
 Db ||||||| : : : ||| ||||||| ||||||| ||||||| ||||||| |||||||  
 QY 1510 CCACCGAAGCCAGTTCAAAGGCCAGGAGCGGATCTTTGGGAAACTGAAAGAGGAAAC 1569  
 Db ||||||| : : : ||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 QY 437 ProProGluSerGlnPheLysAlaGlnGlyArgIlePheGlyLysLeuLysGluGluAsn 456  
 Db ||||||| : : : ||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 QY 1570 TTCTTTAACCCCAAGAGAGTGAAGTGAAGCGCATATCAGAGTGCCCTCTCCACCA 1629  
 Db ||||||| : : : ||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 QY 457 PhePheAsnProLysGluAspValLysLeuGluThrHisIleArgValProSerSerThr 476  
 Db ||||||| : : : ||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 QY 1630 GCTGGCGGGTGATTGCAAGGCGCAAGACCGTGAACGACTGACGAGACTTAAACAGT 1689  
 Db ||||||| : : : ||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 QY 477 AlaGlyArgValIleGlyLysGlyLysThrValAsnGluLeuGluAsnLeuIleSer 496  
 Db ||||||| : : : ||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 QY 1690 GCAGAAGTCACTCGTCCTGTGACCAACGCCAGATGAAATAGAGGAAGTGATCGTCAGA 1749  
 Db ||||||| : : : ||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 QY 497 AlaGluValIleValProArgAspGlnThrProAspGluAsnGluMetIleValArg 516  
 Db ||||||| : : : ||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 QY 1750 ATTATCGGGCACTCTTTGCTAGCCAGACTGCACAGCGCAAGATCAGGGAAATGTACAA 1809  
 Db ||||||| : : : ||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 QY 517 IleIleGlyHisPheAlaSerGlnThrAlaGlnArgLysIleArgGluIleValGln 536  
 Db ||||||| : : : ||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 QY 1810 CAGGTGAAGCAGCAGGAGCAGAAATACCTCAGGAGTCCGCTCACAGCGCAGCAAG 1866  
 Db ||||||| : : : ||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
 QY 537 GlnValLysGlnGlnGlnLysTyProGlnGlyValAlaSerGlnArgSerLys 555

## RESULT 7

AAY30649  
 ID AAY30649 standard; protein; 577 AA.

AC AAY30649;

XX DT 17-NOV-1999 (first entry)  
 XX DE A murine c-myc coding region determinant binding protein.

XX KW c-myc coding region determinant binding protein; CRD-BP; tumor; c-myc;

XX KW endonucleolytic attack; half-life; breast cancer; colon cancer;  
 XX KW pancreatic cancer.

XX OS Mus musculus.

XX PN WO9946594-A2.

XX PD 16-SEP-1999.

XX PF 05-MAR-1999; 99WO-US004897.

XX PR 09-MAR-1998; 98US-0077372P.

XX PA (WISC ) WISCONSIN ALUMNI RES FOUND.

XX PI Ross J;

XX DR WPI; 1999-551506/46.

XX DR N-PSDB; AAZ10617.

XX DR Diagnosing presence or absence of a tumor in a human by examining c-myc  
 XX PT coding region determinant-binding protein.

XX PS Example; Fig 1A-D; 79pp; English.

XX CC The present sequence represents a murine c-myc coding region determinant  
 CC binding protein (CRD-BP). The presence or absence of a tumor can be  
 CC determined by determining the levels of CRD-BP present in the suspect  
 CC tissue, where the CRD-BP shields c-myc RNA from endonucleolytic attack  
 CC and so prolongs its half-life. The methods are used for diagnosing  
 CC presence or absence of a tumor in a human, especially breast, colon and  
 CC pancreatic cancer. They are also used to inhibit cancer cell growth

XX SQ Sequence 577 AA;

Alignment Scores:  
 Pred. No.: 1.47e-173 Length: 577  
 Score: 2019.50 Matches: 394  
 Percent Similarity: 78.94% Conservative: 82  
 Best Local Similarity: 65.34% Mismatches: 96  
 Query Match: 32.99% Indels: 31  
 DB: 2 Gaps: 4

US-09-270-437D-6 (1-3412) x AAY30649 (1-577)

QY 73 ATGACAAAGCTTTACATCGGGAACCTCAGCCCGCTCAGCCCGGACGACCTCCGCGAG 132  
 Db 1 MetAsnLysLeuTyrlleGlyAsnLeuAsnGluSerValThrProAlaLeuGluLys 20  
 QY 133 CTCTTTGGGACAGGAAGCTGCCCTCGCGGACAGGTCCTGCTGAAGTCGGGTACGCC 192  
 Db 21 ValPheAlaGluHisIleSerTyrlleSerGlyGlnPheLeuValLysSerGlyTyrlleAla 40  
 QY 193 TTCTGTGACTACCCGACACAGAACTGGGCCATCCGGCCATCGAGACCTCTCGGGTAAA 252  
 Db 41 PheValAspCysProAspGluHisTrpAlaMetLysAlaIleGluThrPheSerGlyLys 60  
 QY 253 GTGCAATTGCATGGGAAATCATGTGAAGTTGATTCTCAGTCTCTAAAACCTTAAGCAGC 312  
 Db 61 ValGluLeuGlnGlyLysArgLeuGluMetGluHisSerValProLysLysGlnArgSer 80  
 QY 313 AGGAAATTCAGATTTCGAAACATCCCTCTCAGCTCAGTGGGAGGTGTTGGATGCACTT 372  
 Db 81 ArgLysIleGlnIleArgAsnIleProGlnLeuArgTrpGluValLeuAspSerLeu 100  
 QY 373 TTGCTCAATATGGGACAGTGGAGAATGTGGAACAAGTCAACACAGACACAGAAACCGCC 432  
 Db 101 LeuAlaGlnTyrlleGlyThrValGluAsnGluGlnValAsnThrGluSerGluThrAla 120  
 QY 433 GTTGTCAACGTCAATATGCAACAGAGAGAGCAAAAATAGCCATGCGAGAGCTTAAGC 492  
 Db 121 ValValAsnValThrTyrlleSerAsnArgGluGlnThrArgGlnAlaIleMetLysLeuAsn 140  
 QY 493 GGGCATCAGTTTGAAGAACTACTCTTCAAGATTTCCTACATCCCGGATGAAGAGTGAAGC 552  
 Db 141 GlyHisGlnLeuGluAsnHisAlaLeuLysValSerTyrlleProAspGluGlnIleThr 160  
 QY 553 TCCCTTTCGCCCTCAGCGAGCCCGAGCGTGGGACCACTCTTCCCGGGAGAGGCCAC 612  
 Db 161 -----GlnGlyProGluAsnGlyArgArgGlyGlyPheGlySerArgGlyGlnProArg 178  
 QY 613 -----GCCCTGGGGGCACTTCTCAGGCCAGACAGATTGATTTCCCGCTG 657  
 Db 179 GlnGlySerProValAlaAlaGlyAlaProAlaLysGlnGlnProValAspIleProLeu 198  
 QY 658 CGGATCCTGTCCCGACCCAGTTTGTGTGGTCCCATCATCGAAAGAGGGGCTTGACCATTA 717  
 Db 199 ArgLeuLeuValProThrGlnTyrlleValGlyAlaIleIleGlyLysGluGlyAlaThrIle 218  
 QY 718 AAGAACATCACTAAGCAGACCCAGTCCCGGGTGAATATCCATAGAAAAGAGACTCTGGA 777  
 Db 219 ArgAsnIleThrLysGlnThrGlnSerLysIleAspValHisArgLysGluAsnAlaGly 238  
 QY 778 GCTGCGAGAGAGCTGTCAACATCCATGCCACCCAGAGGGGACTTCTGAAGCATGCCG 837  
 Db 239 AlaAlaGluLysAlaIleSerValHisSerThrProGluGlyCysSerSerAlaCysLys 258  
 QY 838 ATGATTCTTGAATATCATGCAAGAGGCGAGATGACACCAAACTAGCCGAGAGATTCTCT 897  
 Db 259 MetIleLeuGluIleMetHisLysGluAlaLysAspThrLysThrAlaAspGluValPro 278  
 QY 898 CTGAAATCTTGGCACACAACTGCTGGTTGGAAGCTGATTGGAAGAAAGAGCAGAAAT 957  
 Db 279 LeuLysIleLeuAlaHisAsnAsnPheValGlyArgLeuIleGlyLysGluGlyArgAsn 298  
 QY 958 TTCAAGAAATTCGAACATGAACAGGAGCCACAGATAACATCTCATCTTTGAGGATTTC 1017

299	LeuLysLysValGluGlnAspThrGluThrLysIleThrIleSerSerLeuGlnAspLeu	318
1018	AGCATATACAAACCCGGAAAGAACCATACACTGTGAAGGGCCACAGTTGAGGCTGTGCCAGT	1077
319	ThrLeuTyrAsnProGluArgThrIleThrValLysGlyAlaIleGluAsnCysCysArg	338
1078	GCTGAGATAGAGATTATGAAGAAGCTCGGTGAGGCGCTTTGAAATGATATGCTGGCTGTT	1137
339	AlaGluGlnGluIleMetLysValArgGluAlaTyrGluAsnAspValAlaAlaMet	358
1138	AACCAACAAGCAACTCTGATCCCAAGGTTGAACCTCAGCGCACCTTGTCATCTTTTCAACA	1197
359	SerLeuGlnSerHisLeuIleProGlyLeuAsnLeuAlaValGlyLeuPheProAla	378
1198	GGACTGTGCGTGTATCTCCACAGCAGGCGCCCGCGAGAGCTCCCCCGCTGCCCTCATC	1257
379	SerSerSerAlaValProProPro-----ProSerSerValThrGlyAlaAlaProTyr	396
1258	CACCCCTTCACTACCCACTCCGGATACTTCTCCAGCGCTGTACCCCATCCAGTTTGGC	1317
397	SerSerPheMet-----	400
1318	CCGTTCCCGCATCATCTCTTTATCCAGAGCAGAGATGTGTAATCTCTCATCCCAACC	1377
401	-----GlnAlaProGluGlnGluMetValGlnValPheIleProAla	414
1378	CAGCTGTGGCGCCATCATCGGAAGAAGGGGCACACATCAACAGCTGGCGAGATTC	1437
415	GlnAlaValGlyAlaIleIleGlyLysGlyGlnHisIleLysGlnLeuSerArgPhe	434
1438	GCGGAGCCTCTATCAAGATTGCCCTCGCGAAGCCACACGTCACGCAAAAGGATGGTC	1497
435	AlaSerAlaSerIleLysIleAlaProGluThrProAspSerLysValArgMetVal	454
1498	ATCATCACCGGCCACCGGAGCCAGTTCAGGCCCCAGGACGGATCTTTGGAAACTG	1557
455	ValIleThrGlyProProGluAlaGlnPheLysAlaGlnGlyArgIleTyrGlyLysLeu	474
1558	AAACAGGAAAACTCTTTAAACCCCAAGAGAAGTGAAGCTGGAAGCGCATATCAGAGTG	1617
475	LysGluGluAsnPhePheGlyProLysGluValLysLeuGluThrHisIleArgVal	494
1618	CCCTCTCCACAGCTGGCGCGGTGATTTGGCAAGGTCGGAGACCGTGAACAGCTCCAG	1677
495	ProAlaSerAlaAlaGlyArgValIleGlyLysGlyLysThrValAsnGluLeuGln	514
1678	AACTTAACCAAGTCAGAGAAGTCATCGTCCTCGTGACCAAAAGCCAGATGAAATGAGAA	1737
515	AsnLeuThrAlaAlaGluValValProArgAspGlnThrProAspGluAsnAspGln	534
1738	GTGATCGTCAGAAATTATCGGCATCTCTTTGCTAGCCAGACTGCACAGCGCAAGATCAGG	1797
535	ValIleValLysIleIleGlyHisPheTyrAlaSerGlnMetAlaGlnArgLysIleArg	554
1798	GAAATTGTACAAACAGGTGAACACACAGCAGAGAGAATACCCCTCAGGAGTGCCTCACAG	1857
555	AspIleLeuAlaGlnValLysGlnGlnHisGlnLysGlyGlnSerAsnLeuAlaGlnAla	574
1858	CGCAGCAAG	1866
575	ArgArgLys	577

DECIIT.T 8

RESULT 8  
AAB11365

AA11365  
ID AA11365 standard: protein: 579 AA.

XX  
XX  
COSTITUTY DIAA  
AC AAB11365:

AC  
XX  
DT

DI	ZI-FEB-2001 (first entry)
XX	
DE	Human lung cancer associated antigen L523S.

DE human lung cancer associated antigen L3238.  
XX  
XW Lung cancer; therapy; treatment; human; tumor; immunogenic; cytostatic; human lung cancer associated antigen L3238.

KW	vaccine; detection.
XX	
XX	Homo sapiens.
XX	
XX	W02000061612-A2.
XX	
XX	19-OCT-2000.
XX	
XX	03-APR-2000; 2000WO-US008896.
XX	
XX	02-APR-1999; 99US-00285479.
XX	17-DEC-1999; 99US-00466396.
XX	30-DEC-1999; 99US-00476496.
XX	10-JAN-2000; 2000US-00480884.
XX	22-FEB-2000; 2000US-00510376.
XX	
XX	(CORI-) CORIXA CORP.
XX	
XX	Wang T, Fan L;
XX	
XX	WPI; 2000-628399/60.
XX	N-PSDB; AAC66035.
XX	
XX	Isolated polypeptide comprising an immunogenic portion of a lung tumor
XX	protein is used for detecting and monitoring progression of lung cancer
XX	in a patient.
XX	
XX	Claim 3; Page 259-261; 261pp; English.
XX	
XX	This invention describes a novel isolated polypeptide (I) which
XX	comprising an immunogenic portion of a lung tumor protein or variant (P2)
XX	which have cytostatic activity. The polypeptides and polynucleotides are
XX	used in compositions and vaccines to inhibit the development of cancer,
XX	especially lung cancer, in a patient. Methods described in the invention
XX	can be used to monitor the progression of a cancer by carrying out the
XX	detection at subsequent time points and comparing the results from the
XX	different time points. CD4+ and/or CD8+ T-Cells isolated from a patient
XX	are treated with P2, polynucleotides encoding P2 or antigen presenting
XX	cells expressing P2 and then administered to the patient to inhibit
XX	development of cancer
XX	
XX	Sequence 579 AA;
XX	
XX	Sequence 579 AA;
XX	

Db 81 ArgLysLeuGlnIleArgAsnIleProProHisLeuGlnTrpGluValLeuAspSerLeu 100



CC comprising the lung tumour proteins, polynucleotides, antibodies, fusion  
CC proteins, T cell populations, or antigen presenting cells that express  
CC the lung tumour proteins are useful for treating lung cancer or  
CC stimulating an immune response. ABL48959 to ABL49300 and ABB74946 to  
CC ABB75070 represent sequences used in the exemplification of the present  
CC invention

XX SQ Sequence 579 AA;

## Alignment Scores:

Pred. No.:	6,18e-168	Length:	579
Score:	1957.50	Matches:	387
Percent Similarity:	79.19%	Conservative:	81
Best Local Similarity:	65.48%	Mismatches:	94
Query Match:	31.98%	Indels:	29
DB:	5	Gaps:	8

US-09-270-437D-6 (1-3412) x ABB75053 (1-579)

QY	73	ATGAACAAGCTTTATCATCGGAACCTGAGCCCGCGCTCACCGCGAGCACTCCCGGCAG	132
DB	1	MetAsnLysLeuTyrIleGlyAsnLeuSerGluAsnAlaAlaProSerAspLeuGluSer	20
QY	133	CTCTTTGGGGACAGAGCTGCCCTGGCGGACAGAGTCTCTGTAAGTCCGGCTACGCC	192
DB	21	IlePheLysAspAlaLysIleProValSerGlyProPheLeuValLysThrGlyTyrAla	40
QY	193	TTCTGTGAGTACCCCGACCAAGTGGCGCATCGCGCCATCGAGACCTCTCGGGTAA	252
DB	41	PheValAspCysProAspGluSerTrpAlaLeuLysAlaIleGluAlaLeuSerGlyLys	60
QY	253	GTGGAATTGCTGGGAAATCATGGAAGTTGATTACTCAGTCTCTAAAAGCTAAGGAGC	312
DB	61	IleGluLeuHisGlyLysProIleGluValGluHisSerValProLysArgGlnArgIle	80
QY	313	AGGAAATTGATTCGAAACATCCCTCTCCTCAGTGGAGGTCTTGATGGACTT	372
DB	81	ArgLysLeuGlnIleArgAsnIleProProHisLeuGlnTrpGluValLeuAspSerLeu	100
QY	373	TTGGCTCAATATGGACAGTGGAGAATGTGAAAGTCAACAGACACACAGAACCGCC	432
DB	101	LeuValGlnTyrGlyValValGluSerCysGluGlnValAsnThrAspSerGluThrAla	120
QY	433	GTTCACACGTCACATATGCAACAGAGAGCAAAATAGCCATGGAGAAGCTAAGC	492
DB	121	ValValAsnValThrTyrSerSerLysAspGlnAlaArgGlnAlaLeuAspLysLeuAsn	140
QY	493	GGGCATCAGTTTGAACTACTCTTCAAGATTTCCTACATCCCGGATGAAGAGTGAGC	552
DB	141	GlyPheGlnLeuGluAsnPheThrLeuLysValAlaTyrIleProAspGluThrAlaAla	160
QY	553	TCCCTTTCGCCCCCTCAG-----CGAGCCACAGCT-----GGGACCACTCTTCCCGG	600
DB	161	GlnGlnAsnProLeuGlnGlnProArgGlyArgArgGlyLeuGlyGlnArgGlySerSer	180
QY	601	GAGCAAGCCACGCCCTGGGGCACTTCTCAGGCCACAGAGATTGATTTCCCGCTGGG	660
DB	181	ArgGlnGly---SerProGlySerValSerLysGlnLysProCysAspLeuProLeuArg	199
QY	661	ATCTGTGTCACCCAGTTTGTGTCCTCATCGGAAGAGGAGGCTTGACCATAAAG	720
DB	200	LeuLeuValProThrGlnPheValGlyAlaIleIleGlyLysGluGlyAlaThrIleArg	219
QY	721	AACATCACTAAGACAGACCCAGTCCGGGTAGATATCCATGAAAGAGAACTCTGGAGCT	780
DB	220	AsnIleThrLysGlnThrGlnSerLysIleAspValHisArgLysGluAsnAlaGlyAla	239
QY	781	GCAGAGAGCCTGTCATCCTCATTCATGCCACCCAGAGGGAATCTGAGCATGCCGATG	840
DB	240	AlaGluLysSerIleThrIleLeuSerThrProGluGlyThrSerAlaAlaCysLysSer	259
QY	841	ATTCCTTGAATCATGCAGAAGAGCGAGATGAGACCAACTAGCCGAGAGATTCCTCTG	900

RESULT 10

ABB74997

ID ABB74997 standard; protein; 579 AA.

XX

AC ABB74997;

DT 01-MAY-2002 (first entry)

XX	Human lung tumour L523S protein sequence SEQ ID NO:348.	
DE	Human; lung tumour; lung cancer; cytostatic; immunostimulant; vaccine; immune response.	
XX	Homo sapiens.	
XX	WO200200174-A2.	
XX	03-JAN-2002.	
XX	28-JUN-2001; 2001WO-US021065.	
XX	28-JUN-2000; 2000US-00606421.	
PR	02-AUG-2000; 2000US-00630940.	
PR	21-AUG-2000; 2000US-00643597.	
PR	15-SEP-2000; 2000US-00662786.	
PR	09-OCT-2000; 2000US-00685696.	
PR	12-DEC-2000; 2000US-00735705.	
PR	07-MAY-2001; 2001US-00850716.	
XX	(CORI-) CORIXA CORP.	
XX	Wang T, Wang A, Skeiky YAW, Li SX, Kalos MD, Henderson RA; McNeill PD, Fanger N, Retter MW, Warnerakis M, Fanger GR; Vedvick TS, Carter D, Watanabe Y, Peckham DW;	
XX	WPI; 2002-090513/12.	
DR	N-PSDB; ABL49254.	
XX	Polynucleotides encoding lung tumor polypeptides, useful for treating lung cancer or stimulating an immune response.	
XX	Example 2; Page 330-332; 374pp; English.	
XX	The present invention describes human lung tumour proteins. Human lung tumour proteins and polynucleotides have cytostatic and immunostimulant activities, and can be used in vaccine production. Compositions comprising the lung tumour proteins, polynucleotides, antibodies, fusion proteins, T cell populations, or antigen presenting cells that express the lung tumour proteins are useful for treating lung cancer or stimulating an immune response. ABL48959 to ABL49300 and ABB74946 to ABB75070 represent sequences used in the exemplification of the present invention	
XX	Sequence 579 AA;	
XX	Alignment Scores:	
Pred. No.:	6.18e-168	Length: 579
Score:	1957.50	Matches: 387
Percent Similarity:	79.19%	Conservative: 81
Best Local Similarity:	65.48%	Mismatches: 94
Query Match:	31.98%	Indels: 29
DB:	5	Gaps: 8
US-09-270-437D-6 (1-3412) x ABB74997 (1-579)		
QY	73 ATGAACAAGCTTTACATCGGGAACCTGAGCCCGCCCGCTACCGCCGACGACCTCCGGCAG	132
DB	1 MetAsnLysLeuTyrlleGlyAsnLeuSerGluAsnAlaProSerAspLeuGluSer	20
QY	133 CTCTTTGGGACAGAGCTGCCCTCGCGGACAGGTCCTGCTGAGTCGGGTAGCCGCTAGCC	192
DB	21 IlePheLysAspAlaLyslleProValSerGlyProPheLeuValLysThrGlyTyra	40
QY	193 TTCGTGGACTACCCGACCAAGTGGGCGCATCGCGCCATCGAGACCTCTCGGTA	252
DB	41 PheValAspCysProAspGlySerlrPAlaLeuLysAlalleGluAlaLeuSerGlyLys	60
QY	253 GTGGAATTGATGGGAAATCATGGAATTGATTACTACTGCTCTAAAGACTAAGGAGC	312
DB	61 IleGluLeuHisGlyLysProIleGluValGluHisSerValProLysArgGlnArgIle	80
QY	313 AGGAAAATTCAGATTCTGAAACATCCCTCCTCAGCTGAGTGGAGGTGGTGGTACGACIT	372
DB	81 ArgLysLeuGlnleArgAsnleProProHisLeuGlnTrpGluValleuAspSerLeu	100
QY	373 TTGGCTCAATATATGGGACAGTGGAGAATGTGGAACAAGTCAACACACACAGAAACCGCC	432
DB	101 LeuValGlnTyGlyValValGluSerCysGluGlnValasnThrAspSerGluThrAla	120
QY	433 GTTGTCAACGTCAATATCAACAGAGAAAGCAAAATAGCCATGGAGAGCTAAGC	492
DB	121 ValValasnValThrTySerSerLysAspGlnAlaArgGlnAlaLeuAspLysLeuAsn	140
QY	493 GGGCATCAGTTTGGAGAACTACTCTTCAAGATTTCAGATCCCGATCCGATGAGAGGTGAGC	552
DB	141 GlyPheGlnGluAsnPheThrLeuLysValAlaTyrlleProAspGluThrAlaAla	160
QY	553 TCCCTTCGCCCCCTCAG-----CGAGCCACGCGT-----GGGACCCACTCTTCCCGG	600
DB	161 GlnGlnAsnProLeuGlnGlnProArgGlyArgArgGlyLeuGlyGlnArgGlySerSer	180
QY	601 GAGCAAGCCACGCCCTCGGGGCACTTCTCAGCCGACAGACAGATTGATTTCCGCTGCGG	660
DB	181 ArgGlnGly---SerProGlySerValSerLysGlnLysProCysAspLeuProLeuArg	199
QY	661 ATCTCGTCCACCCAGTTTGTGGTGCATCATCGGAAGGAGGCTTGACATAAAG	720
DB	200 LeuLeuValProThrGlnPheValGlyAlaIlelleGlyLysGluGlyAlaThrIleArg	219
QY	721 AACATCACTAAGCAGACCCAGTCCCGGTAGATATCATAGAAAAGAACTCTGGAGCT	780
DB	220 AsnIleThrLysGlnThrGlnSerLyslleAspValHisArgLysGluAsnAlaGlyAla	239
QY	781 GCAGAGAAGCTGTCCATCCATCCATGCCACCCAGAGGGGACTTCTGAAGATCCCGCATG	840
DB	240 AlaGluLysSerIleThrIleLeuSerThrProGluGlyThrSerAlaAlaCysLysSer	259
QY	841 ATTCTTGAATCATGCAAGAGGCGAGATGAGACCACTAGCCGAGAGATTCTCTCTG	900
DB	260 IleLeuGluIleMetHisLysGluAlaGlnAspIleLysPheThrGluGluIleProLeu	279
QY	901 AAAATCTTGGCACACAATGGCTTGGTGAAGACTGATTGAAAAGAGGAGCAGAAATTG	960
DB	280 LysIleLeuAlaHisasnAsnPheValGlyArgLeuIleGlyLysGluGlyArgAsnLeu	299
QY	961 AAGAAAATGAACATGAACAGGGAACAAGATACCAATCTCATCTTTTCAGAGATTTCAGC	1020
DB	300 LysLysIleGluGlnAspThrAspThrLysIleThrIleSerProLeuGlnGluLeuThr	319
QY	1021 ATATACACCCGGAAGAACCATCACTGTGAAGGGCACAGTTGAGGCTCTGCGAGTCT	1080
DB	320 LeuTyAsnProGluArgThrIleThrValLysGlyAsnValGluThrCysAlaLysAla	339
QY	1081 GAGATAGAGATTATGAAGAGCTGCGTGAAGCCCTTTGAAAATGATATGCTGGCTGTTAAC	1140
DB	340 GluGluGluIleMetLysLyslleArgGluSerTyGluAsnAspIleAlaSerMetAsn	359
QY	1141 CAACAGCCAAATCTGATCCAGGGTGAACCTCAGCGCACTTGGCATCTTTTTCACAGGA	1200
DB	360 LeuGlnAlaHisLeuIleProGlyLeuAsnLeuAsnAlaLeuGlyLeuPhe-----	376
QY	1201 CTGTCGCTGTATCTCCACAGGCGGCGGAGCTCCCGCGCTCCCGCTGCGCCCTACAC	1260
DB	377 -----ProThrSer-----GlyMetProProProThr-----	386
QY	1261 CCCTTCACTACCCACTCCGGATACTTCTCCAGCTGTACCCCATCACCAGTTTGGCCCG	1320
DB	387 -----SerGlyProProSerAlaMetThrPro-----Pro	396
QY	1321 TTCCCGCATCATCCTCTTATCCAGAGCAGGAGATTGTGAATCTCTTATCCCAACCCAG	1380
DB	397 TyrProGlnPheGluGln---SerGluThrGluThrValHisLeuPheIleProAlaLeu	415





Db 240 AladgluysSerIleThrIleuSerThrProGluGlyThrSerAlaAlaCysLysSer 259  
 Qy 841 ATTCTTGAATCATGACAGAGCAGATGAGACCAAACTAGCGAAGAGATTCCTGTG 900  
 Db 260 ILeuGluIleMetHisLysGluAlaGlnAspLysPheThrGluGluIleProLeu 279  
 Qy 901 AAAATCTTGGCACACAATGGCTGTGGTGGAGACTGATTGGAAGAGAGGAGCAAAATTG 960  
 Db 280 LysIleLeuAlaHisAsnAsnPheValGlyArgLeuIleGlyLysGluGlyArgAsnLeu 299  
 Qy 961 AAGAAAATTGAACATGAACAGGACCAAGATACAACTCTCATCTTTTCGACGATTGAGC 1020  
 Db 300 LysLysIleGluGlnAspThrAspThrLysIleThrIleSerProLeuGlnGluLeuThr 319  
 Qy 1021 ATATACACCCGGAAGAACCATCATCTGTGAGGGCAGCTGAGGCGCTGTGCCAGTGCT 1080  
 Db 320 LeuTyrAsnProGluArgThrIleThrValLysGlyAsnValGluThrCysAlaLysAla 339  
 Qy 1081 GAGATAGAGATTATGAAGAAGCTCGTGGAGGCTTGAACCTCAGCGCACTTGGCATCTTTTCAACAGGA 1200  
 Db 360 LeuGlnAlaHisLeuIleProGlyLeuAsnLeuAlaLeuGlyLeuPhe----- 376  
 Qy 1201 CTGTCCGTGTATCTCCACGAGAGGCGCCCGGAGGCTCCCGCGTGCCTTACAC 1260  
 Db 377 -----ProThrSer-----GlyMetProProThr----- 386  
 Qy 1261 CCCTTCACTCCCACTCGGATCTTCTCCAGCTGTACCCCGCATCCAGCTTGGCGCG 1320  
 Db 387 -----SerGlyProProSerAlaMetThrPro-----Pro 396  
 Qy 1321 TTCCCGCATCATCTCTTATCCAGCAGCAGAGATTGTGAATCTCTTTCATCCCAACCCAG 1380  
 Db 397 TyrProGlnPheGluGln---SerGluThrGluThrValHisLeuPheIleProAlaLeu 415  
 Qy 1381 GCTGTGGCGGCATCATCGGGAAGAGAGGGGCAACATCAACAGCTGGCGAGATTGCGC 1440  
 Db 416 SerValGlyAlaIleIleGlyLysGlnGlyGlnHisIleLysGlnLeuSerArgPheAla 435  
 Qy 1441 GGAGCTCTATCAGATTGCCCTCGGAGAGGCGCCAGAGCTGACGAGAAAGATCGTCTATC 1500  
 Db 436 GlyAlaSerIleLysIleAlaProAlaGluAlaProAspAlaLysValArgMetValIle 455  
 Qy 1501 ATCACCAGGCGCCAGGAGCCAGCTTCAAGGCGCCAGGAGCGATCTTTGGGAAACTGAAA 1560  
 Db 456 IleThrGlyProProGluAlaGlnPheLysAlaGlnGlyArgIleTyrGlyLysIleLys 475  
 Qy 1561 GAGAAAACCTTCTTTAACCCCAAGAGAACTGAGCTGGAAGCGCATATCAGAGTGCC 1620  
 Db 476 GluGluAsnPheValSerProLysGluGluValLysLeuGluAlaHisIleArgValPro 495  
 Qy 1621 TCTTCCACAGCTGCGCGGTGATTGGCAAGAGTGGCAAGCCGTGAACGACTCGAGAC 1680  
 Db 496 SerPheAlaIleGlyArgValIleGlyLysGlyLysThrValAsnGluLeuGlnAsn 515  
 Qy 1681 TTAACCAAGTCAGAGATCATCGTCCCTCGTGACCAAAACCCAGATGAAATAGAGAAAGTG 1740  
 Db 516 LeuSerSerAlaGluValValProArgAspGlnThrProAspGluAsnAspGlnVal 535  
 Qy 1741 ATCGTCAGATTATCGGGCACTCTTCTGTAGCCAGACTGACAGCGCAAGATCAGGAA 1800  
 Db 536 ValValLysIleThrGlyHisPheTyrAlaCysGlnValAlaGlnArgLysIleGlnGlu 555  
 Qy 1801 ATTGTACAAAGGTGAAGCAGCAGAGCAGAA 1833  
 Db 556 IleLeuThrGlnValLysGlnHisGlnGlnGln 566

RESULT 12  
 ABP61917

ID ABP61917 standard; protein; 579 AA.  
 AC ABP61917;  
 DT 07-OCT-2002 (first entry)  
 XX Human lung cancer associated protein sequence SEQ ID NO:348.  
 DE Human; lung cancer; lung tumour; cytostatic; gene therapy; vaccine.  
 KW Homo sapiens.  
 XX WO200247534-A2.  
 PN 20-JUN-2002.  
 PD 30-NOV-2001; 2001WO-US047576.  
 XX 12-DEC-2000; 2000US-00735705.  
 PR 07-MAY-2001; 2001US-00850716.  
 PR 28-JUN-2001; 2001US-00897778.  
 XX (CORI-) CORIXA CORP.  
 PA Wang T, Wang A, Skeiky YAW, Li SX, Kalos MD, Henderson RA;  
 PI Mcneill PD, Fanger N, Retter MW, Durham M, Fanger GR, Vedwick TS;  
 PI Carter D, Watanabe Y, Peckham DW, Cai F, Foy TM;  
 DR WPI; 2002-583465/62.  
 DR N-PSDB; ABQ92440.  
 XX Novel lung carcinoma polynucleotide sequences and polypeptides encoded by  
 PT the polynucleotides, useful in pharmaceutical compositions such as  
 PT vaccines and as markers to indicate the presence of lung cancer.  
 XX Example 2; Page 337-339; 381pp; English.  
 PS The present invention describes isolated human lung carcinoma  
 CC polynucleotides (I) and polypeptides (II). (I) and (II) have cytostatic  
 CC activity, and can be used in gene therapy and in vaccines. Compositions  
 CC comprising (I) or (II) can be used for stimulating an immune response in  
 CC a patient and for treating lung cancer in a patient. Oligonucleotides of  
 CC (I) can be used for detecting the presence of a cancer in a patient, by  
 CC obtaining a biological sample from the patient, contacting the biological  
 CC sample with the oligonucleotide, detecting in the sample, an amount of  
 CC polynucleotide that hybridises to the oligonucleotide and comparing the  
 CC amount of polynucleotide that hybridises to the oligonucleotide to a  
 CC predetermined cut-off value, and determining the presence of a cancer in  
 CC the patient. (I) and (II) are useful in pharmaceutical compositions, e.g.  
 CC vaccines. (I) is useful as a marker to indicate the presence or absence  
 CC of a cancer such as lung cancer. ABQ92145 to ABQ92486 and ABP61866 to  
 CC ABP61992 represent sequences used in the exemplification of the present  
 CC invention  
 SQ Sequence 579 AA;  
 Alignment Scores:  
 Pred. No.: 6.18e-168 Length: 579  
 Score: 1957.50 Matches: 387  
 Percent Similarity: 79.19% Conservative: 81  
 Best Local Similarity: 65.48% Mismatches: 94  
 Query Match: 31.98% Indels: 29  
 DB: 5 Gaps: 8  
 US-09-270-437D-6 (1-3412) x ABP61917 (1-579)  
 Qy 73 ATGACAGCTTTACATCGGAGACCTGAGCCCGCGCTACCGCCGACGACCTCCGCGAG 132  
 Db 1 MetAsnLysLeuTyrIleGlyAsnLeuSerGluAsnAlaAlaProSerAspLeuGluSer 20  
 Qy 133 CTCTTTGGGACAGGAGCTGCCCTGGCGGACAGCTCTCTGCTGAGTCCGCGCTAGCC 192  
 Db 21 IlePheLysAspAlaLysIleProValSerGlyProPheLeuValLysThrGlyTyrAla 40

QY	193	TTCTGGGACTACCCGACAGAACTGGGCGCATCCGCGCATCGAGACCTCTCGGATAA	252
Db	41	PheValAspCysProAspGluSerTrpAlaLeuLysAlaIleGluAlaLeuSerGlyLys	60
QY	253	GTGGAAATTCATGCGGAAATCATGGAAGTTGATTACTCAGTCTCTAAAAGCTAAGGAGC	312
Db	61	IleGluLeuHisGlyLysProIleGluValGluHisSerValProLysArgGlnArgIle	80
QY	313	AGGAAATTCAGATTGCAAAACATCCCTCTCCTCAGTGCAGTGGAGGTGTGATGACTT	372
Db	81	ArgLysLeuGlnIleArgAsnIleProProHisLeuGlnTrpGluValLeuAspSerLeu	100
QY	373	TTGGCTCAATATGGGACAGTGGAGAATGTGGAACAAGTCAACACAGACACAGAAACCGCC	432
Db	101	LeuValGlnTyrGlyValValGluSerCysGluGlnValAsnThrAspSerGluThrAla	120
QY	433	GTGTCAACGTCACATATCAACAAGAGAAAGCAAAATAGCCATGCGAGAGCTAAGC	492
Db	121	ValValAsnValThrTyrSerSerLysAspGlnAlaArgGlnAlaLeuAspLysLeuAsn	140
QY	493	GGGCATCATGTTGAGAACTACTCTTCAGATTCTACATCCGATCCGATCAAGGTGAGC	552
Db	141	GlyPheGlnLeuGluAsnPheThrLeuLysValAlaTyrIleProAspGluThrAlaAla	160
QY	553	TCCCTCTGCCCCCTCAG-----CGAGCCACAGCGT-----GGGGACCACTCTTCCCGG	600
Db	161	GlnGlnAsnProLeuGlnGlnProArgGlyArgArgGlyLeuGlyGlnArgGlySerSer	180
QY	601	GAGCAAGGCCAGCCCTCGGGGCATCTCTCAGGCCAGACAGATTGATTTCCGCTGGG	660
Db	181	ArgGlnGly---SerProGlySerValSerLysGlnLysProCysAspLeuProLeuArg	199
QY	661	ATCTGTGTCCTCCACCCAGTTTGTGTGTCATCATCGGAAGAGGGCTTGACCATAAAG	720
Db	200	LeuLeuValProThrGlnPheValGlyAlaIleGlyLysGluGlyAlaThrIleArg	219
QY	721	AACATCATCAAGCAGACCCAGTCCCGGTAGATATCCATAGAAAAGAGAACTCTCGAGCT	780
Db	220	AsnIleThrLysGlnThrGlnSerLysIleAspValHisArgLysGluAsnAlaGlyAla	239
QY	781	GCAGAGACCTGTACCATCCATGCCATGCCACCCAGAGGGGACTTCTGAAGCATGCCGATG	840
Db	240	AlaGluLysSerIleThrIleLeuSerThrProGluGlyThrSerAlaAlaCysLysSer	259
QY	841	ATTCTTGAATCATGCAGAGAGGCAGATGAGACCAACTAGCCGACAGATTCTCTG	900
Db	260	IleLeuGluIleMethHisLysGluAlaGlnAspIleLysPheThrGluGluIleProLeu	279
QY	901	AAAACTTTGGCCACACAATGGCTTGGTTGGAAGACTGATTGGAAAAGAGCCAGAAATTG	960
Db	280	LysIleLeuAlaHisAsnAsnPheValGlyArgLeuIleGlyLysGluGlyArgAsnLeu	299
QY	961	AGAAATTTGAATCAATGAACAGGACCAAGATAACAATCTCATCTTTGAGGATTGAGC	1020
Db	300	LysLysIleGluGlnAspThrAspThrLysIleThrIleSerProLeuGlnGluLeuThr	319
QY	1021	ATATACACCCGGAAGACCATCACTGTGAAGGGCAGAGTTGAGCCCTGTCGAGTGCT	1080
Db	320	LeuTyrAsnProGluArgThrIleThrValLysGlyAsnValGluThrCysAlaLysAla	339
QY	1081	GAGATAGATTATGAAGAGCTGCGTGAAGGCTTTGAAATGATATGCTGTTTAAAC	1140
Db	340	GluGluGluIleMethLysLysIleArgGluSerTyrGluAsnAspIleAlaSerMetAsn	359
QY	1141	CAACAAGCAATCTGATCCAGGGTTGAACCTCAGCGGCACTTGGCATCTTTCAACAGA	1200
Db	360	LeuGlnAlaHisLeuIleProGlyLeuAsnLeuAsnAlaLeuGlyLeuPhe-----	376
QY	1201	CTGTCCGTGCTATCTCACCAGCAGGGCCCCGCGAGCTCCCGCTGCCCCCTACCAC	1260
Db	377	-----ProProThrSer-----GlyMetProProThr-----	386
QY	1261	CCCTTCACTACCCACTCCGGATPACTTCTCCAGCTGTATCCCCCATCACCCAGTTTGGCCCG	1320
Db	387	-----SerGlyProProSerAlaMetThrPro-----Pro	396
QY	1321	TTCCGCGATCATCACTCTTATCCAGAGCAGAGATTGGAATCTTCTATCCCAACCAG	1380
Db	397	TyrProGlnPheGluGln---SerGluThrGluThrValHisLeuPheIleProAlaLeu	415
QY	1381	GCTGTGGCGCCATCATCGGAAGAAGGGGCGACACATCAACACAGCTGCGAGATTGCGC	1440
Db	416	SerValGlyAlaIleIleGlyLysGlnGlnHisIleLysGlnLeuSerArgPheAla	435
QY	1441	GGAGCTCTATCAAGATTGCCCTCGGAAGCCAGAGCTCAGCGAAAAGGATGTGTCATC	1500
Db	436	GlyAlaSerIleLysIleAlaProAlaGluAlaProAspAlaLysValArgMetValIle	455
QY	1501	ATCACCGGCGCACCCGGAAGCCAGTTCAGAGCCAGGAGCGGATCTTTGGGAAACTGAAA	1560
Db	456	IleThrGlyProProGluAlaGlnPheLysAlaGlnGlyArgIleTyrGlyLysIleLys	475
QY	1561	GAGGAAAATCTTCTTTAACCCCAAGAAAGTGAAGCTGGAAGCGCATATCAGAGTGCC	1620
Db	476	GluGluAsnPheValSerProLysGluGluValLysLeuGluAlaHisIleArgValPro	495
QY	1621	TCTTCCACAGCTGGCGGGTGATTGGCAAGAGTGCAAGACCGTGAAACGAACTGCAGAAC	1680
Db	496	SerPheAlaAlaGlyArgValIleGlyLysGlyLysThrValAsnGluLeuGlnAsn	515
QY	1681	TTAACCACTGCAGAAAGTATCGTCTGTGACCAAAAGCCAGATGAAAATCAGGAAGTG	1740
Db	516	LeuSerSerAlaGluValValProArgAspGlnThrProAspGluAsnAspGlnVal	535
QY	1741	ATCGTCAGAAATTTATCGGCACCTTCTTGTCTACCCAGACTGCACAGCGCAAGATCAGGAA	1800
Db	536	ValValLysIleThrGlyHisPheLysAlaCysGlnValAlaGlnArgLysIleGlnGlu	555
QY	1801	ATTGTACAACTGAGTGAAGCAGCAGGAGCAGAAA	1833
Db	556	IleLeuThrGlnValLysGlnHisGlnGlnGln	566
RESULT 13			
ABP61974			
ID	ABP61974	standard; protein; 579 AA.	
XX	XX	ABP61974;	
XX	XX	07-OCT-2002 (first entry)	
DT	DT	Human lung cancer associated protein sequence SEQ ID NO:449.	
DE	DE	Human; lung cancer; lung tumour; cytostatic; gene therapy; vaccine.	
XX	XX	Homo sapiens.	
OS	OS	WO200247534-A2.	
XX	XX	20-JUN-2002.	
XX	XX	30-NOV-2001; 2001WO-US047576.	
PF	PF	12-DEC-2000; 2000US-00735705.	
PR	PR	07-MAY-2001; 2001US-00850716.	
PR	PR	28-JUN-2001; 2001US-00897778.	
XX	XX	(CORI-) CORIXA CORP.	
PA	PA	Wang T, Wang A, Skeiky YAW, Li SX, Kalos MD, Henderson RA;	
XX	XX	McNeill PD, Fanger N, Retter MW, Durham M, Fanger GR, Vedvick TS;	
PI	PI	Carter D, Watanabe Y, Peckham DW, Cai F, Foy TW;	
XX	XX	WPI: 2002-583465/62.	
DR	DR	N-PSDB; ABQ92485.	
XX	XX		

PT Novel lung carcinoma polynucleotide sequences and polypeptides encoded by  
PT the polynucleotides, useful in pharmaceutical compositions such as  
PT vaccines and as markers to indicate the presence of lung cancer.

PS Claim 9; Page 375-377; 381pp; English.

The present invention describes isolated human lung carcinoma polynucleotides (I) and polypeptides (II). (I) and (II) have cytostatic activity, and can be used in gene therapy and in vaccines. Compositions comprising (I) or (II) can be used for stimulating an immune response in a patient and for treating lung cancer in a patient. Oligonucleotides of (I) can be used for detecting the presence of a cancer in a patient, by obtaining a biological sample from the patient, contacting the biological sample with the oligonucleotide, detecting in the sample, an amount of polynucleotide that hybridises to the oligonucleotide and comparing the amount of polynucleotide that hybridises to the oligonucleotide to a predetermined cut-off value, and determining the presence of a cancer in the patient. (I) and (II) are useful in pharmaceutical compositions, e.g. vaccines. (I) is useful as a marker to indicate the presence or absence of a cancer such as lung cancer. ABQ92145 to ABQ92486 and ABP61866 to ABP61992 represent sequences used in the exemplification of the present invention

Sequence 579 AA;

Alignment Scores:

Pred. No.:	6.18e-168	Length:	579
Score:	1957.50	Matches:	387
Percent Similarity:	79.19%	Conservative:	81
Best Local Similarity:	65.48%	Mismatches:	94
Query Match:	31.98%	Indels:	29
DB:	5	Gaps:	8

US-09-270-437D-6 (1-3412) x ABP61974 (1-579)

QY	73	ATGAACAAGCTTTACATCGGGAACCTGAGCCAGCCCGCGGTACCGCCGACACGACTCTCGGCGAG	132
Db	1	MetAsnLysLeuTyrIleGlyAsnLeuSerGluAsnAlaAlaProSerAspLeuGluSer	20
QY	133	CTCTTTGGGGACAGAAAGCTGCCCTCGCGGACACAGGCTCTGCTGAAGTCCGGCTACGCG	192
Db	21	IlePheLysAspAlaLysIleProValSerGlyProPheLeuValLysThrGlyTyrAla	40
QY	193	TTCTGTGGACTACCCCGACACAACTGGGCATCCCGCCATCGAGACCTCTCGGGTAA	252
Db	41	PheValAspCysProAspGluSerThrPalaLeuLysAlaIleGluAlaLeuSerGlyLys	60
QY	253	GTGGAAATTCATGGGAAATCATGGAAGTTGATTACTCAGTCTCTAAAAGCTAAAGGAC	312
Db	61	IleGluLeuHisGlyLysProIleGluValGluHisSerValProLysArgGlnArgIle	80
QY	313	AGGAAATTCAGTTCGAAACATCCCTCTCACTGTCAGTGGAGAGTGTGTGATGAGCTT	372
Db	81	ArgLysLeuGlnIleArgAsnIleProProHisLeuGlnTyrGluValLeuAspSerLeu	100
QY	373	TTGGCTCAATATGGACAGTGGAGAAATGTGGACAAAGTCAACACAGACACAGAACCCGC	432
Db	101	LeuValGlnTyrGlyValIleGluSerCysGluGlnValAsnThrAspSerGluThrAla	120
QY	433	GTTGTCAACGTCACATATGCACAGAGAAGCAAAANTAGCCATGCAGAGACTAAGC	492
Db	121	ValValAsnValThrLysSerLysAspGlnAlaArgGlnAlaLeuAspLysLeuAsn	140
QY	493	GGGCATCATGTTTGAACTACTCTCTCAAGATTCTCATATCCCGATCAAGAGGTGAC	552
Db	141	GlyPheGlnLeuLysAsnPheThrLeuLysValAlaTyrIleProAspGluThrAlaAla	160
QY	553	TCCCTCTCGCCCTTCAG-----CGAGCCCAAGCT-----GGGACCACTCTTCCCGG	600
Db	161	GlnGlnAsnProLeuGlnGlnProArgGlyArgGlyLeuGlyClnArgGlySerSer	180
QY	601	GAGCAAGGCCACGCCCTTGGGGGCATCTCTCAGGCCACAGATTGATTTCCGCTGCGG	660

QY	1741	ATCGTCAGAAATATCGGCGACTTCTTTGCTAGCCAGACTGCACAGCGCAAGATCAGGAA	1800
		ValVallylsileThrGlyHisPheTyrAlaCysGlnValAlaGlnArgLysIleGlnGlu	555
Db	536	ATTCGTACAAACAGGTGAAGCAGCAGGAGCAGAAA	1833
QY	1801	ATTCGTACAAACAGGTGAAGCAGCAGGAGCAGAAA	1833
		IleLeuThrGlnVallylsGlnHisGlnGlnGln	566
Db	556	IleLeuThrGlnVallylsGlnHisGlnGlnGln	566
RESULT 14			
ABP61973			
ID	ABP61973	standard; protein; 579 AA.	
XX	AC	ABP61973;	
XX			
XX		07-OCT-2002 (first entry)	
XX			
DE		Human lung cancer associated protein sequence SEQ ID NO:446.	
XX			
KW		Human; lung cancer; lung tumour; cytostatic; gene therapy; vaccine.	
XX			
OS		Homo sapiens.	
XX			
PN		WO200247534-A2.	
XX			
PD		20-JUN-2002.	
XX			
PF		30-NOV-2001; 2001WO-US047576.	
XX			
PR		12-DEC-2000; 2000US-00735705.	
PR		07-MAY-2001; 2001US-00850716.	
PR		28-JUN-2001; 2001US-00897778.	
XX			
PA		(CORI-) CORIXA CORP.	
XX			
PI		Wang T, Wang A, Skeiky YAW, Li SX, Kalos MD, Henderson RA;	
PI		Mcneill PD, Fanger N, Retter WM, Durham M, Fanger GR, Vedwick TS;	
PI		Carter D, Watanabe Y, Peckham DW, Cai F, Foy TM;	
XX			
DR		WPI; 2002-583465/62.	
DR		N-PSDB; ABQ92483.	
XX			
PT		Novel lung carcinoma polynucleotide sequences and polypeptides encoded by	
PT		the polynucleotides, useful in pharmaceutical compositions such as	
PT		vaccines and as markers to indicate the presence of lung cancer.	
XX			
PS		Claim 9; Page 372-374; 381pp; English.	

DB: 5 Caps: 8

US-09-270-437D-6 (1-3412) x ABP61973 (1-579)

QY 73 ATGAACAAGCTTTACATCGGAACTGAGCCCGCGCTCACGCCGACGACCTCCGGCAG 132  
 |||||  
 Db 1 MetAsnLysLeuTyrIleGlyAsnLeuSerGluAsnAlaAlaProSerAspLeuGluSer 20  
 |||||  
 QY 133 CTCTTTGGGACAGGAAAGCTGCCCTGGCGGGAGCAGGTCTCTGTAAGTCCGGCTACGCC 192  
 |||||  
 Db 21 IlePheLysAspAlaIysIleProValSerGlyProPheLeuValLysThrGlyTyrAla 40  
 |||||  
 QY 193 TTCGTGAGCTACCCCGACAGAACTCGGCCATCCGCGCCATCGAGACCTCTCGGATAA 252  
 |||||  
 Db 41 PheValAspCysProAspGluSerThrPalaLeuLysAlaIleGluAlaLeuSerGlyLys 60  
 |||||  
 QY 253 GTGAATTGCATCGGAAAATCATGGAAGTTGATTACTCAGTCTCTAAAAGCTTAAGGAGC 312  
 |||||  
 Db 61 IleGluLeuHisGlyLysProIleGluValGluHisSerValProLysArgGlnArgIle 80  
 |||||  
 QY 313 AGGAAATTCAGATTCGAAACATCCCTCTCACTCAGTGGAGGTGTGGATGCACATT 372  
 |||||  
 Db 81 ArgLysLeuGlnIleArgAsnIleProProHisLeuGlnTrpGluValLeuAspSerLeu 100  
 |||||  
 QY 373 TTGGCTCAATATCGGACAGTGGAGAATGTGGAACAAGTCAACACACACAGAAACCGCC 432  
 |||||  
 Db 101 LeuValGlnTyrGlyValValGluSerCysGluGlnValAsnThrAspSerGluThrAla 120  
 |||||  
 QY 433 GTTGCTCACTCATATGCAACAGAGAAGAAGCAAAATAGCCATGGAGAGCTAAGC 492  
 |||||  
 Db 121 ValValAsnValThrTyrSerLysAspGlnAlaArgGlnAlaLeuAspLysLeuAsn 140  
 |||||  
 QY 493 GGGCATCAGTTTCAGAACTACTCCTTCAGATTTCCTACATCCCGATGAAGAGGTGAGC 552  
 |||||  
 Db 141 GlyPheGlnLeuGluAsnPheThrLeuLysValAlaTyrIleProAspGluThrAlaAla 160  
 |||||  
 QY 553 TCOCCTTCGCCCTCAG-----CGAGCCAGCGT-----GGGACCACTTCTCCCG 600  
 |||||  
 Db 161 GlnGluAsnProLeuGlnGlnProArgGlyArgArgGlyLeuGlyGlnArgGlySerSer 180  
 |||||  
 QY 601 GAGCAGGCCACCCCTCGGGGCACTTCTCAGGCCAGACAGATTGATTTCCCGTCCGG 660  
 |||||  
 Db 181 ArgGlnGly---SerProGlySerValSerLysGlnLysProCysAspLeuProLeuArg 199  
 |||||  
 QY 661 ATCTGTGTCCTCCACCCAGCTTTGTGTGTCATCATCGAAAGAGAGCGCTTGACCTAAAG 720  
 |||||  
 Db 200 LeuLeuValProThrGlnPheValGlyAlaIleIleGlyLysGluGlyAlaThrIleArg 219  
 |||||  
 QY 721 AACTCATCTAAGCAGACCCCGTCCGGTAGATATCCATAGAAAAGAGAACTCTGGAGCT 780  
 |||||  
 Db 220 AsnIleThrLysGlnThrGlnSerLysIleAspValHisArgLysGluAsnAlaGlyAla 239  
 |||||  
 QY 781 GCAGAGAAGCTCTCACCATCCATCCACCCACAGAGGGGACTTCTGAAGCTGCCCGATG 840  
 |||||  
 Db 240 AlaGluLysSerIleThrIleLeuSerThrProGluGlyThrSerAlaAlaCysLysSer 259  
 |||||  
 QY 841 ATTCTTGAATCATGCAGAAAGAGGCGAGATGAGACCAAACTAGCCGAGAGATTCCTCTG 900  
 |||||  
 Db 260 IleLeuGluIleMethHisLysGluAlaGlnAspIleLysPheThrGluGluIleProLeu 279  
 |||||  
 QY 901 AAAATCTTGCCACACAATGGCTTGTGTGGAAGACTGATTGGAAAAAGAGCGCAAAATTG 960  
 |||||  
 Db 280 LysIleLeuAlaHisAsnAsnPheValGlyArgLeuIleGlyLysGluArgLysLeuLeu 299  
 |||||  
 QY 961 AAGAAAATTGAAATCAAGAAAGAGGACCAAGATAACAATCTCATCTTTTGAGGATTTGAGC 1020  
 |||||  
 Db 300 LysLysIleGluLysAspThrLysIleThrIleSerProLeuGlnGluLeuThr 319  
 |||||  
 QY 1021 ATATACAACCCGAAAGAACCATCACTGTGAGGGGCACAGTTGAGCCCTGTCAGTGCT 1080  
 |||||  
 Db 320 LeuTyrAsnProGluArgThrIleThrValLysGlyAsnValGluThrCysAlaLysAla 339  
 |||||  
 QY 1081 GAGATAGAGATTATGAAGAAGCTCGCTGAGGCCCTTTGAAAATCATATGCTGCTGCTAAC 1140  
 |||||



```

Db      81  ArgLysLeuGlnIleArgAsnIleProHisLeuGlnTipGluValLeuAspSerLeu 100
      373  TTGGCTCATATGGGACAGTGGAGATGGACAAAGTCAACACACACACACAAACCGCC 432
Db     101  LeuValGlnTyrGlyValGluSerCysGluGlnValAsnThrAspSerGluThrAla 120
      433  GTGTCAACGTCACATATCAACACAGAGAAAGCAAAATAGCCATGAGAGCTTAAGC 492
Db     121  ValValAsnValThrTyrSerSerLysAspGlnAlaArgGlnAlaLeuAspLysLeuAsn 140
      493  GGGCATCATGTTGAGAACTATCTTCAGATTTCTACATCCCGATGAAGAGGTGAGC 552
Db     141  GlyPheGlnLeuGluAsnPheThrLeuLysValAlaTyrIleProAspGluThrAlaAla 160
      553  TCCCTTCGCCCCCTCAG-----CGAGCCAGCGT-----GGGACCACTCTTCCCGG 600
Db     161  GlnGlnAsnProLeuGlnProArgGlyArgGlyLeuGlyGlnArgGlySerSer 180
      601  GAGCAAGGCCACGCCCTGGGGGCACCTCTCAGGCCACACAGATTGATTCCTCGGTGGG 660
Db     181  ArgGlnGly---SerProGlySerValSerLysGlnLysProCysAspLeuProLeuArg 199
      661  ATCCTGTTCCCAACCGAGTTTGTGGTCCCATCATCGGAAAGAGGGCTTGACCATAAAG 720
Db     200  LeuLeuValProThrGlnPheValGlyAlaIleGlyLysGluGlyAlaThrIleArg 219
      721  AACATCACTAAGCACACCCAGTCCCGGTAGATATCCATAGAAAGAGAACTCTCGAGCT 780
Db     220  AsnIleThrLysGlnThrGlnSerLysIleAspValHisArgLysGluAsnAlaGlyAla 239
      781  GCAGAAAGCCCTGCACCATCATGCCACCCAGAGGGGACTTCTGAAGCATGCCGCATG 840
Db     240  AlaGluLysSerIleThrIleLeuSerThrProGluGlyThrSerAlaAlaCysLysSer 259
      841  ATTCTTGAATCATCCAGAGAGGACAGATGAGACCAACTAGCCAGAGATTCTCTCTG 900
Db     260  IleLeuGluIleMetHisLysGluAlaGlnAspIleLysPheThrGluGluIleProLeu 279
      901  AAAATCTTGGCACAAATGGCTTGTGGAGACTGATTGGAAAGAGAGGACAGAAATTTG 960
Db     280  LysIleLeuAlaHisAsnAsnPheValGlyArgLeuIleGlyLysGluGlyArgAsnLeu 299
      961  AAGAAATTTGACATGAACAGGACCAAGGACCAAGATTAACAATCTCATCTTTGAGGATTGAGC 1020
Db     300  LysLysIleGluGlnAspThrAspThrLysIleThrIleSerProLeuGlnGluLeuThr 319
      1021  ATATACACCCGGAAGAACCATCACTGTGAAGGACACAGTTGAGGCCTGTGCCAGTGT 1080
Db     320  LeuTyrAsnProGluArgThrIleThrValLysGlyAsnValGluThrCysAlaLysAla 339
      1081  GAGATAGAGATTATGAAGAAGCTGCGTAGGGCTTTGAAATGATATGCTGCTGTTAAC 1140
Db     340  GluGluGluIleMetLysLysIleArgGluSerTyrGluAsnAspIleAlaSerMetAsn 359
      1141  CAACAAAGCAATCTGATCCCGGGTTGAACCTCAGCGCATTTGGCATCTTTTCAACAGGA 1200
Db     360  LeuGlnAlaHisLeuIleProGlyLeuAsnLeuAsnAlaLeuGlyLeuPhe----- 376
      1201  CTGTCCGTGCTATCTCCACACGAGGCGCCCGGAGCTCCCGCGCTGCCCGCTTACCAC 1260
Db     377  -----ProProThrSer-----GlyMetProProThr----- 386
      1261  CCCTTCACTACCCATCCGGATACCTTCTCCAGCCCTGTACCCCATCAGGTTTGGCCCG 1320
Db     387  -----SerGlyProProSerAlaMetThrPro-----Pro 396
      1321  TTCCCGCATCATCACTCTTATCCAGAGCAGAGAGATTGGAATCTTCTATCCCAACCCAG 1380
Db     397  TyrProGlnPheGluGln---SerGluThrThrValHisLeuPheIleProAlaLeu 415
      1381  GCTGTGGCGCCCATCATCGGGAAGAGGGGGCACATCAACAGCTGGCGAGATTGCGCC 1440
      ::::|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

```

Db     416  SerValGlyAlaIleIleGlyLysGlnGlyGlnHisIleLysGlnLeuSerArgPheAla 435
      1441  CGAGCCTCTATCAAGATTGCCCTCGCGAAGCCCGAGAGCTCAGCGAAGGATGTCATC 1500
Db     436  GlyAlaSerIleLysIleAlaProAlaGluAlaProAspAlaLysValArgMetValIle 455
      1501  ATCACCAGGCGCCACCGAAGCCAGTTCAAGGCCAGGCGGACCGGATCTTTGGGAAACTGAAA 1560
Db     456  IleThrGlyProProGluAlaGlnPheLysAlaGlnGlyArgIleTyrGlyLysIleLys 475
      1561  GAGGAAATCTTCTTTAAACCCCAAGAAAGTGAAGTGGAGCCCATATCAGAGTGGCC 1620
Db     476  GluGluAsnPheValSerProLysGluGluValLysLeuGluAlaHisIleArgValPro 495
      1621  TCTTCCACAGCTGGCGGGTGTGTCGCAAGGTGGCAAGCCGTCGAACGAAGTGCAGAAC 1680
Db     496  SerPheAlaAlaGlyArgValIleGlyLysGlyGlyLysThrValAsnGluLeuGlnAsn 515
      1681  TTAACCAAGTCAGAAAGTCACTGCTGCTGACCAAAAGCCAGATGAAATGAGGAAGTG 1740
Db     516  LeuSerSerAlaGluValValValProArgAspGlnThrProAspGluAsnAspGlnVal 535
      1741  ATCGTCAGAAATATCGGGCACTTCTTGTAGCCAGACTGCACAGCGCAGATCAGGAA 1800
Db     536  ValValLysIleThrGlyHisPheTyrAlaCysGlnValAlaGlnArgLysIleGlnGlu 555
      1801  ATTGTACAACAGGTGAAGCAGCAGGACGAGCAAAA 1833
      556  IleLeuThrGlnValLysGlnHisGlnGlnGln 566

```

Search completed: July 16, 2004, 10:38:50  
Job time : 191 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: July 16, 2004, 10:30:59 ; Search time 33 Seconds  
(without alignments)  
10675.625 Million cell updates/sec

Title: US-09-270-437D-6  
Perfect score: 6121  
Sequence: 1 ggcagcgaggagcgaggagga.....aacctgaaatgtttattt 3412

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 778828

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ n2p.model -DEV=xlp  
-Q=/cgn2\_1/USPTO.spool\_p/US09270437/runat 16072004 113050 13370/app query.fasta\_1.3591  
-DB=Issued Patents AA -OPMT=fastan -SUFFIX=rai -MINMATCH=0 1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09270437 @CGN 1.1 46 @runat 16072004 113050 13370 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents AA:\*

1: /cgn2\_6/ptodata/2/iaa/5A.COMB.pep.\*  
2: /cgn2\_6/ptodata/2/iaa/5B.COMB.pep.\*  
3: /cgn2\_6/ptodata/2/iaa/6A.COMB.pep.\*  
4: /cgn2\_6/ptodata/2/iaa/6B.COMB.pep.\*  
5: /cgn2\_6/ptodata/2/iaa/PTCUS.COMB.pep.\*  
6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2019.5	33.0	577	3	US-09-261-855-2
2	1957.5	32.0	579	4	US-09-643-597-348
3	1957.5	32.0	579	4	US-09-542-615A-348
4	1957.5	32.0	579	4	US-09-606-421B-348
5	1950.5	31.9	579	4	US-09-643-597-176
6	1950.5	31.9	579	4	US-09-480-884A-176
7	1950.5	31.9	579	4	US-09-542-615A-176
8	1950.5	31.9	579	4	US-09-606-421B-176
9	245	4.0	644	1	US-08-021-608D-2
10	245	4.0	644	1	US-08-726-160-2
11	245	4.0	644	5	PCT-US94-01782-2
12	243.5	4.0	643	1	US-08-021-608D-10

13	243.5	4.0	643	1	US-08-726-160-10	Sequence 10, Appl
14	243.5	4.0	643	5	PCT-US94-01782-10	Sequence 10, Appl
15	236.5	3.9	590	1	US-08-021-608D-8	Sequence 8, Appl
16	236.5	3.9	590	1	US-08-726-160-8	Sequence 8, Appl
17	236.5	3.9	590	5	PCT-US94-01782-8	Sequence 8, Appl
18	226	3.7	530	1	US-08-187-793-4	Sequence 4, Appl
19	215	3.5	48	3	US-09-261-855-20	Sequence 20, Appl
20	210	3.4	48	3	US-09-261-855-24	Sequence 24, Appl
21	210	3.4	49	3	US-09-261-855-18	Sequence 18, Appl
22	205	3.3	720	4	US-09-252-991A-21881	Sequence 21881, A
23	201	3.3	47	3	US-09-261-855-21	Sequence 21, Appl
24	200	3.3	49	3	US-09-261-855-22	Sequence 22, Appl
25	197	3.2	47	3	US-09-261-855-17	Sequence 17, Appl
26	195	3.2	1706	4	US-09-252-991A-31760	Sequence 31760, A
27	190	3.1	1020	4	US-09-252-991A-28970	Sequence 28970, A
28	186.5	3.0	1476	4	US-09-252-991A-29427	Sequence 29427, A
29	184.5	3.0	1228	4	US-09-252-991A-17764	Sequence 17764, A
30	184	3.0	343	1	US-08-187-793-2	Sequence 2, Appl
31	181	3.0	748	4	US-09-252-991A-18427	Sequence 18427, A
32	177	2.9	47	3	US-09-261-855-23	Sequence 23, Appl
33	176.5	2.9	471	3	US-08-866-928B-1	Sequence 1, Appl
34	176.5	2.9	471	4	US-09-685-836-1	Sequence 1, Appl
35	176	2.9	47	3	US-09-261-855-19	Sequence 19, Appl
36	176	2.9	705	4	US-09-252-991A-30792	Sequence 30792, A
37	175.5	2.9	1002	4	US-09-252-991A-27980	Sequence 27980, A
38	174.5	2.9	1780	1	US-08-769-309A-5	Sequence 5, Appl
39	174.5	2.9	1780	3	US-08-994-570-5	Sequence 5, Appl
40	174	2.8	657	4	US-09-252-991A-28001	Sequence 28001, A
41	173.5	2.8	705	4	US-09-252-991A-23298	Sequence 23298, A
42	172	2.8	719	4	US-09-252-991A-23660	Sequence 23660, A
43	171.5	2.8	697	4	US-09-252-991A-24009	Sequence 24009, A
44	169.5	2.8	2294	4	US-09-252-991A-17231	Sequence 17231, A
45	169	2.8	742	4	US-09-252-991A-32659	Sequence 32659, A

ALIGNMENTS

RESULT 1  
US-09-261-855-2  
; Sequence 2, Application US/09261855A  
; Patent No. 6255055  
; GENERAL INFORMATION:  
; APPLICANT: Ross, Jeffrey  
; TITLE OF INVENTION: THE C-MYC CODING REGION DETERMINANT-BINDING PROTEIN  
; FILE REFERENCE: (CRD-BP) AND ITS NUCLEIC ACID SEQUENCE  
; FILE REFERENCE: 960296.95131  
; CURRENT APPLICATION NUMBER: US/09/261,855A  
; CURRENT FILING DATE: 1999-03-03  
; NUMBER OF SEQ ID NOS: 46  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 2  
; LENGTH: 577  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-09-261-855-2

Alignment Scores:					
Pred. No.:	1.24e-170	Length:	577		
Score:	2019.50	Matches:	394		
Percent Similarity:	78.94%	Conservative:	82		
Best Local Similarity:	65.34%	Mismatches:	96		
Query Match:	32.99%	Indels:	31		
DB:	3	Gaps:	4		
US-09-270-437D-6 (1-3412) x US-09-261-855-2 (1-577)					
Qy	73	ATGACAGCTTACATCGGAGACCTGAGCCGCCCTACCGCCGACGACCTCGGCAG	132		
Db	1	MetAsnLysLeuTyrIleGlyAsnLysLeuSerValThrProAlaAspLeuGluLys	20		
Qy	133	CTCTTGGGACAGGAGCTGCCCTGGCGGACAGCTCTGCTGAAGTCCGCTAGCC	192		
Db	21	ValPheAlaGluHisIleSerTyrSerGlycInPheLeuValLysSerGlyTyrAla	40		

```
QY 193 TTCGTGGACTACCCGACAGAACTGGGCGCATCCGCGCATCGAGACCTCTCGGTAAA 252
Db 41 PheValAspCysProAspGluHisTrpAlaMetLysAlaIleGluThrPheSerGlyLys 60
QY 253 GTGGAATTGCATGGGAAATCATGGAAGTTGATTACTCAGTCTCTAAAAGCTAAGGAGC 312
Db 61 ValGluLeuGlnGlyLysArgLeuGluMetGluHisSerValProLysLysGlnArgSer 80
QY 313 AGGAAATTCAGATTCGAAACATCCCTCTCTACCTGCGAGTGGAGGTGTGGATGGACTT 372
Db 81 ArgLysIleGlnIleArgAsnIleProProGlnLeuArgTrpGluValLeuAspSerLeu 100
QY 373 TTGGCTCAATATGGGACAGTGGAGATGGGAACAGTCAACACAGACACAGAAACCGCC 432
Db 101 LeuAlaGlnTyrGlyThrValGluAsnCysGluGlnValAsnThrGluSerGluThrAla 120
QY 433 GTTGTCAACGTCACATATGCAACAAGAGAAAGCAAAATAGCCATGAGAGCTAAGC 492
Db 121 ValValAsnValThrTyrSerAsnArgGluGlnThrArgGlnAlaIleMetLysLeuAsn 140
QY 493 GGGCATCATGTTGAACTACTCTTCAGATTTCCTACATCCCGGATCAAGAGTGGAGC 552
Db 141 GlyHisGlnLeuGluAsnHisAlaLeuLysValSerTyrIleProAspGluGlnIleThr 160
QY 553 TCCCTTCGCCCCCTCAGCGAGCCCGCGTGGGGACCACTCTTCCGGGAGCAAGGCCAC 612
Db 161 -----GlnGlyProGluAsnGlyArgGlyGlyPheGlySerArgGlyGlnProArg 178
QY 613 -----GCCCTGGGGCACTTCTCAGGCCGACAGAGATTGATTTCCCGCTG 657
Db 179 GlnGlySerProValAlaAlaGlyAlaProAlaLysGlnGlnProValAspIleProLeu 198
QY 658 CGGATCCGTGTCACCCAGTTTGTGTGGTCCATCATCGGAAGAGGGCTTGACCAT 717
Db 199 ArgLeuLeuValProThrGlnTyrValGlyAlaIleIleGlyLysGluGlyAlaThrIle 218
QY 718 AAGACATCATCAAGCAGACCCAGTCCCGGTPAGATATCCATAGAAAGAGAACTCTGGA 777
Db 219 ArgAsnIleThrLysGlnThrGlnSerLysIleAspValHisArgLysGluAsnAlaGly 238
QY 778 GCTGAGAGAGCCGTGTACCATCCATGCCACCCAGAGGGGACTTCTGAAGCATGCCGC 837
Db 239 AlaAlaGluLysAlaIleSerValHisSerThrProGluGlyCysSerSerAlaCysLys 258
QY 838 ATGATCTTGAATCATCAGAAAGAGGAGCAGATGAGACCACTAGCCGCGAGATTCTCT 897
Db 259 MetIleLeuGluIleMetHisLysGluAlaLysAspThrLysThrAlaAspGluValPro 278
QY 898 CTGAAATCTTGGCACACAATGGCTTGGTTGGAAGACTGATTGGAAGAAAGAGCGAAGAT 957
Db 279 LeuLysIleLeuAlaHisAsnAsnPheValGlyArgLeuIleGlyLysGluGlyArgAsn 298
QY 958 TTGAAGAAATTTGAATCAACACAGGACCAAGATAACAATCTCATCTTTGTCAGGATTG 1017
Db 299 LeuLysLysValGluAlaAspThrGluThrLysIleThrIleSerSerLeuGlnAspLeu 318
QY 1018 AGCATATACACCCGGAAGACCATCACTGTGAAGGAGCAGAGTTGAGSCCTGTGCCAGT 1077
Db 319 ThrLeuTyrAsnProGluArgThrIleThrValLysGlyAlaIleGluAsnCysCysArg 338
QY 1078 GCTGAGATGAGATTATGAAGAAGCTGGGTGAGGCTTTGAAATGATATGCTGCTGTT 1137
Db 339 AlaGluGlnGluIleMetLysLysValArgGluAlaTyrGluAsnAspValAlaAlaMet 358
QY 1138 AACCAACAGCAATCTGATCCAGGTTGAACTCAGGCGACTTGGCATCTTTTCAACA 1197
Db 359 SerLeuGlnSerHisIleProGlyLeuAsnLeuAlaAlaValGlyLeuPheProAla 378
QY 1198 GGACTGTCTGCTATCTCCACGACGAGGGCCCGGAGCTCCCGCGCTCCCGCTAC 1257
Db 379 SerSerSerAlaValProPro-----ProSerSerValThrGlyAlaAlaProTyr 396

RESULT 2
US-09-643-597-348
; Sequence 348, Application US/09643597
; Patent No. 6426072
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy
; APPLICANT: Fanger, Gary R.
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Henderson, Robert A.
; APPLICANT: McNeill, Patricia D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C11
; CURRENT APPLICATION NUMBER: US/09/643,597
; NUMBER OF SEQ ID NOS: 369
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 348
; LENGTH: 579
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-643-597-348
```



Alignment Scores:

Pred. No.: 4, 02e-165 Length: 579  
Score: 1957.50 Matches: 367  
Percent Similarity: 79.19% Conservative: 81  
Best Local Similarity: 65.48% Mismatches: 94  
Query Match: 31.98% Indels: 29  
DB: 4 Gaps: 8

US-09-270-437D-6 (1-3412) x US-09-643-597-348 (1-579)

```
QY 73 ATGAAACAGCTTACATCGGAACCTGAGCCCGCCGCTACCGCGCAGCACTCCGGCAG 132
Db 1 MetAsnLysLeuTyrIleGlyAsnLeuSerGluAsnAlaAlaProSerAspLeuGluSer 20
QY 133 CTCCTTTGGGACAGGAGCTCCCTCGCGGACAGCTCCCTGCTGAACTCCGGTACGCC 192
Db 21 IlePheLysAspAlaLysIleProValSerGlyProPheLeuValLysThrGlyTyrAla 40
QY 193 TTCGTGGACTACCCGACAGAACTGGGCCATCCGCGCCATCGAGACCTCTCGGTAAA 252
Db 41 PheValAspCysProAspGluSerTrpAlaLeuLysAlaIleGluAlaLeuSerGlyLys 60
QY 253 GTGGAATTCATGGAAATCATGAAGTTGATTAAGTCTCTAAAGACTTAAGAGC 312
Db 61 IleGluLeuHisGlyLysProIleGluValGluHisSerValProLysArgGluArgIle 80
QY 313 AGGAAATTCAGATTGAAACATCCCTCCATCGCTGAGTGGAGGAGTGTGGTGGACTT 372
Db 81 ArgLysLeuGlnIleArgAsnIleProHisLeuGlnTrpGluValLeuAspSerLeu 100
QY 373 TTGCTCAATATGGGACAGTGGAGAATGTGAACAAGTCAACACAGACACAGAAACCGCC 432
Db 101 LeuValGlnTyrGlyValValGluSerCysGluGlnValAsnThrAspSerGluThrAla 120
QY 433 GTTGTCAAGTCATATGCAACAGAGAGAGAACAAATAAGCATCGAGAACTAAGC 492
Db 121 ValValAsnValThrTyrSerSerLysAspGlnAlaArgGlnAlaLeuAspLysLeuAsn 140
QY 493 GGGCATCAGTTTGAGAACTACTCTCAAGATTCTTACATCCCGGATGAGAGGTAGC 552
Db 141 GlyPheGlnLeuGluAsnPheThrLeuLysValAlaTyrIleProAspGluThrAlaAla 160
QY 553 TCCCTCTCGCCCTCAG-----CGAGCCAGCGT-----GGGACCACTCTTCCCGG 600
Db 161 GlnGlnAsnProLeuGlnProArgGlyArgGlyLeuGlyGlnArgGlySerSer 180
QY 601 GAGCAAGCCACGCCCTCGGGGACTTCTCAGCCACAGACAGATTGATTTCCCGCTCGG 660
Db 181 ArgGlnGly---SerProGlySerValSerLysGlnLysProCysAspLeuProLeuArg 199
QY 661 ATCTGTGTCGCCACCCAGTTGTTGTTGGTCCATCATCGGAAGGAGGCTTGACCATAAAG 720
Db 200 LeuLeuValProThrGlnPheValGlyAlaIleIleGlyLysGluGlyAlaThrIleArg 219
QY 721 AACATCATAGAGACCCAGTCCCGGTAGATATCCATAGAAAAGAACTCTGGAGCT 780
Db 220 AsnIleThrLysGlnThrGlnSerLysIleAspValHisArgLysGluAsnAlaGlyAla 239
QY 781 GCAGAGAGCTGTACCATCCATCCATCCACCCGAGGGGACTTCTGAGCATCCCGCATG 840
Db 240 AlaGluLysSerIleThrIleLeuSerThrProGluGlyThrSerAlaAlaCysLysSer 259
QY 841 ATTCTTGAATCATGCAAGAGGAGGATGAGACCAACTAGCCGAGAGATTCCTCTG 900
Db 260 IleLeuGluIleMethIleLysGluAlaGlnAspIleLysPheThrGluGluIleProLeu 279
QY 901 AAATCTTGGCACACATGCTGCTGGTTGGAAGTCAATGGAAGAGGAGGAGGAGGAGGAGG 960
Db 280 LysIleLeuAlaHisAsnAsnPheValGlyArgLeuIleGlyLysGluGlyArgGlnLeu 299
QY 961 AAGAAATTCAGATGAAACAGGAGCAGCAATCAATCTCATCTTTGAGGATTTGAGC 1020
Db 300 LysLysIleGluGlnAspThrAspThrLysIleThrLysIleThrLysIleThrLysIleThr 319
```

```
QY 1021 ATATACAAACCGGAAAGAACCATCTGTGAAGGGCACAGTTGAGGCTGTGCCAGTGT 1080
Db 320 LeuTyrAsnProGluArgThrIleThrValLysGlyAsnValGluThrCysAlaLysAla 339
QY 1081 GAGATAGAGATTATGAAGAAGCTCGGTGAGCCCTTTGAAATGATATATCTGGCTTTAAC 1140
Db 340 GluGluGluIleMetLysLysIleArgGluSerTyrGluAsnAspIleAlaSerMetAsn 359
QY 1141 CAACAAGCAATCTGATCCCAAGGTGAACCTCAGCGCACITGGCATCTTTTCAACAGA 1200
Db 360 LeuGlnAlaHisLeuIleProGlyLeuAsnLeuAlaLeuGlyLeuPhe----- 376
QY 1201 CTGTCGTGTATCTCCACAGAGGGCCCGCGAGCTCCCGCGTGTCCCGCTTACCAC 1260
Db 377 -----ProThrSer-----GlyMetProProProThr----- 386
QY 1261 CCCTTCACTACCACTCGGNATCTTCTCCAGCTGTACCCCATCACCAGTTTGGGCCG 1320
Db 387 -----SerGlyProProSerAlaMetThrPro-----Pro 396
QY 1321 TTCCCGCATCATCTCTATCCAGAGGAGATTGTGAATCTTCTCATCCCAACCCAG 1380
Db 397 TyrProGlnPheGluGln---SerGluThrGluThrValHisLeuPheIleProAlaLeu 415
QY 1381 GCTGTGGCGCCATCATCGGGAAGAGGGGCACACATCAAAACAGCTGGCGAGATTCGCC 1440
Db 416 SerValGlyAlaIleIleGlyLysGlnGlyGlnHisIleLysGlnLeuSerArgPheAla 435
QY 1441 GGAGCCTCTCAGATTGCGGAGTCCCTGCGAAGGCCACACGTCCAGCGAAAGATGGTATC 1500
Db 436 GlyAlaSerIleLysIleAlaProAlaGluAlaProAspAlaLysValArgMetValIle 455
QY 1501 ATCAACCGGCGCACCGGAAGCCAGTTTCAAGCCCGAGGAGCGATCTTTGGGAACTCAA 1560
Db 456 IleThrGlyProProGluAlaGlnPheLysAlaGlnGlyArgIleThrGlyLysIleLys 475
QY 1561 GAGAAACTTCTTTAACCCCAAGAGAGTGAAGCTGGAAGCGCATATCAGAGTGCCTC 1620
Db 476 GluGluAsnPheValSerProLysGluGluValLysLeuGluAlaHisIleArgValPro 495
QY 1621 TCTTCCACAGCTCGCGGTGATTGGCAAGGTGGCAAGCCGTGAACGAACTGCAGAAC 1680
Db 496 SerPheAlaAlaGlyArgValIleGlyLysGlyThrValAsnGluLeuGlnAsn 515
QY 1681 TTAACAGAGTCAGAAAGTCATCGTCCCTCGTGACCAACCGCAGATGAAATGAGGAAGTG 1740
Db 516 LeuSerSerAlaGluValValProArgAspGlnThrProAspGluAsnAspGlnVal 535
QY 1741 ATCGTCAGAAATTATCGGCACTTCTTTGTAGCCAGACTGCACAGCGCAAGATCAGGAA 1800
Db 536 ValValLysIleThrGlyHisPheTyrAlaCysGlnValAlaGlnArgLysIleGlnGlu 555
QY 1801 ATTGTACAACAGGTGAAGCAGCAGGAGCAGAAA 1833
Db 556 IleLeuThrGlnValLysGlnHisGlnGlnGln 566
```

RESULT 3

US-09-542-615A-348  
; Sequence 348, Application US/09542615A  
; Patent No. 6519256  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Tongtong  
; APPLICANT: Fan, Ligu  
; APPLICANT: Kalos, Michael D.  
; APPLICANT: Bangur, Chaitanya S.  
; APPLICANT: Hosken, Nancy A.  
; APPLICANT: Fanger, Gary R.  
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY  
; FILE REFERENCE: 210121.455C8  
; CURRENT APPLICATION NUMBER: US/09/542,615A  
; CURRENT FILING DATE: 2000-04-14

; NUMBER OF SEQ ID NOS: 350  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 348  
; LENGTH: 579  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-542-615A-348

Alignment Scores:  
Pred. No.: 4,02e-165 Length: 579  
Score: 1957.50 Matches: 387  
Percent Similarity: 79.19% Conservative: 81  
Best Local Similarity: 65.48% Mismatches: 94  
Query Match: 31.98% Indels: 29  
DB: 4 Gaps: 8

US-09-270-437D-6 (1-3412) x US-09-542-615A-348 (1-579)

```
QY 73 ATGAACAGCTTTACATCGGAACCTGAGCCCGCGCTCACCGCGACGACCTCCGGCAG 132
Db 1 MetAsnLysLeuTyrlleGlyAsnLeuSerGluAsnAlaA1aProSerAspLeuGluSer 20
QY 133 CTCCTTTGGGGACAGGAAGTCCCTCGGGGACAGGTCCTGCTGAAGTCGCGGTACGCC 192
Db 21 IlePheLysAspAlaLysIleProValSerGlyProPheLeuValLysThrGlyTyrAla 40
QY 193 TTCGTGGACTACCCCGACGACCACTGGCCATCCCGGCCATCGAGACCTCTCCGGTAAA 252
Db 41 PheValAspCysProAspGluSerTrpAlaLeuLysAlaIleGluAlaLeuSerGlyLys 60
QY 253 GTGGAATTGCATGGGAAATCATGGAAGTTGATCTCAGTCTTAAAGAACTAAGAGAGC 312
Db 61 IleGluLeuHisGlyLysProIleGluValGluHisSerValProLysArgGlnArgIle 80
QY 313 AGGAAATTCAGATTGCAAACTCCCTCTACCTGACAGTGGAGGTGTGATGGACTT 372
Db 81 ArgLysLeuGlnIleArgAsnIleProProHisLeuGlnTrpGluValLeuAspSerLeu 100
QY 373 TTGGCTCAATAGGACAGTGGAGATGTGGACAGTCAACACAGACACAGAAACCGCC 432
Db 101 LeuValGlnTyrGlyValValGluSerCysGluGlnValAsnThrAspSerGluThrAla 120
QY 433 GTTGTCAACGTCACATATGCAACAGAGAAAGCAAAATAGCCATGGAGAAGCTTAAGC 492
Db 121 ValValAsnValThrTyrSerSerLysAspGlnAlaArgGlnAlaLeuAspLysLeuAsn 140
QY 493 GGGCATCATGTTGAGAACTATCTTCAAGATTTCCTCATCCCGATGAAGAGTGAGC 552
Db 141 GlyPheGlnLeuGluAsnPheThrLeuLysValAlaTyrIleProAspGluThrAlaAla 160
QY 553 TCCCTCTCCGCCCTCAG-----CGAGCCCGAGCT-----GGGGACCACTCTCCCGG 600
Db 161 GlnGlnAsnProLeuGlnGlnProArgGlyArgGlyLeuGlyGlnArgGlySerSer 180
QY 601 GAGCAGGCGACCGCCCTCGGGGCACTTCTCAGGCACAGACAGATTGATTTCCCGTGGCG 660
Db 181 ArgGlnGly--SerProGlySerValSerLysGlnLysProCysAspLeuProLeuArg 199
QY 661 ATCTGTGTCACCCAGCTTTGTGTGTCATCATCGGAAGAGAGGGCTTGACCATAAAG 720
Db 200 LeuLeuValProThrGlnPheValGlyAlaIleIleGlyLysGluGlyAlaThrIleArg 219
QY 721 AACATCACTAAGCAGACCCAGTCCCGGTAGATATCCATAGAAAAGAGAACTCTGGAGCT 780
Db 220 AsnIleThrLysGlnThrGlnSerLysIleAspValHisArgLysGluAsnAlaGlyAla 239
QY 781 GCAGAGACCTGTACCATTCATCCATCCACCCAGAGGGGACTTCTGAAGCATGCCCATG 840
Db 240 AlaGluLysSerIleThrIleLeuSerThrProGluGlyThrSerAlaAlaCysLysSer 259
QY 841 ATTCTTGAATCATGCAAGAGGCGAGATGAGACCAAACTAGCCCGAGAGATTCTCTCG 900
Db 260 IleLeuGluIleMethHisLysGluAlaGlnAspIleLysPheThrGluGluIleProLeu 279
```

```
QY 901 AAAATCTTGGCACACAATGGCTTGGTTGGAAGACTGATTGGAAGAGGAGGAGAAATTTG 960
Db 280 LysIleLeuAlaHisAsnAsnPheValGlyArgLeuIleGlyLysGluGlyArgAsnLeu 299
QY 961 AAGAAAATTTGAACATGAACAGGACCAAGATAACAATCTCATCTTTCAGAGATTGAGC 1020
Db 300 LysLysIleGluGlnAspThrAspThrLysIleThrIleSerProLeuGlnGluLeuThr 319
QY 1021 ATATACAACCCGGAAGAACCATCACTGTGAAGGCGACAGTTGAGGCTGTGCCAGTCT 1080
Db 320 LeuTyrAsnProGluArgThrIleThrValLysGlyAsnValGluThrCysAlaLysAla 339
QY 1081 GAGATAGAGATTATGAAGAAGCTCGTGAGGCCCTTTGAAAATGATATGCTGCTGTTAAC 1140
Db 340 GluGluGluIleMetLysLysIleArgGluSerTyrGluAsnAspIleAlaSerMetAsn 359
QY 1141 CAACAAGCAATCTGATCCCGAGGTGAACCTCAGCGCAGCTTGGCATCTTTTCAACAGA 1200
Db 360 LeuGlnAlaHisLeuIleProGlyLeuAsnLeuAsnAlaLeuGlyLeuPhe----- 376
QY 1201 CTGTCCGTGCTATCTCCACACGACGAGGCGCCCGCGAGCTCCCGCGCTGCCCTTACC 1260
Db 377 -----ProProThrSer-----GlyMetProProProThr----- 386
QY 1261 CCTTCTACTACCCACTCCGAGATCTTCTCCAGCCTGTACCCCAATCACAGTTTGGCCCG 1320
Db 387 -----SerGlyProProSerAlaMetThrPro-----Pro 396
QY 1321 TTCGCGCATCATCTCTTATCCAGACGAGAGATTGGAATCTCTTATCCCAACCCAG 1380
Db 397 TyrProGlnPheGluGln---SerGluThrGluThrValHisLeuPheIleProAlaLeu 415
QY 1381 GCTGTGGCGGCATCATCGGAAGAGGGGGCACACATCAACAGCTGCGAGATTGCGC 1440
Db 416 SerValGlyAlaIleIleGlyLysGlnGlnHisIleLysGlnLeuSerArgPheAla 435
QY 1441 GGAGCCTCTATCAAGATTGCCCCCTCGGAAGCCCGACGCTCAGCGAAAGGATGTCATC 1500
Db 436 GlyAlaSerIleLysIleAlaProAlaGluAlaProAspAlaLysValArgMetValIle 455
QY 1501 ATCAGCGGCGCCCGAAGCCCGAGTTCAAGGCCAGGACCGGATCTTTGGGAACCTGAA 1560
Db 456 IleThrGlyProProGluAlaGlnPheLysAlaGlnGlyArgIleTyrGlyLysIleLys 475
QY 1561 GAGGAAAATCTCTTTAACCCCAAGAGAGTGAAGCTGGAAGCCGATATCAGAGTGC 1620
Db 476 GluGluAsnPheValSerProLysGluGluValLysLeuGluAlaHisIleArgValPro 495
QY 1621 TCTTCCACAGCTGGCGGGTGTATTCGCAAGGTGCAAGACCGTGAACGAACTGCAGAAC 1680
Db 496 SerPheAlaAlaGlyArgValIleGlyLysGlyLysThrValAsnGluLeuGlnAsn 515
QY 1681 TTAACAGTGCAGAGTATCATCGTGTCTGTGACCAAAAGCCAGATGAAATCAGGAAGT 1740
Db 516 LeuSerSerAlaGluValValProArgAspGlnThrProAspGluAsnGlnVal 535
QY 1741 ATCTGTCAAGATTATCGGCACCTCTTTGTAGCCAGACTGCACAGCGCAAGATCAGGAA 1800
Db 536 ValValLysIleThrGlyHisPheTyrAlaCysGlnValAlaGlnArgLysIleGlnGlu 555
QY 1801 ATTGTACACAGCTGAAGCAGCAGGAGCAGAAA 1833
Db 556 IleLeuThrGlnValLysGlnHisGlnGlnGln 566
```

## RESULT 4

US-09-606-421B-348  
; Sequence 348, Application US/09606421B  
; Patent No. 6531315  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Tongtong  
; APPLICANT: Fan, Ligu  
; APPLICANT: Kalos, Michael D.

; APPLICANT: Bangur, Chaitanya S.		QY		721	AAATCACTAAGCAGACCCAGTCCCGGTAGATATCCATAGAAAAGAACTCTGGAGCT	780
; APPLICANT: Hosken, Nancy		Db		220	AsnIleThrLysGlnThrGlnSerLysIleAspValHisArgLysGluAsnAlaGlyAla	239
; APPLICANT: Fanger, Gary R.		QY		781	GCAGAGAAAGCGTGTCAACCATCCATCCACCCAGAGGGAGTCTTGAAGCATGCCGATG	840
; APPLICANT: Li, Samuel X.		Db		240	AlaGluLysSerIleThrIleLeuSerThrProGluGlyThrSerAlaAlaCysLysSer	259
; APPLICANT: Wang, Aijun		QY		841	ATTCTTGAATCATGTCAGAAAGAGGCGACATGAGACCAAACTAGCCGAGAGATTCTCTG	900
; APPLICANT: Skeiky, Yasir A.W.		Db		260	IleLeuGluIleMetHisLysGluAlaGlnAspIleLysPheThrGluGluIleProLeu	279
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY		QY		901	AAATCTTGGCAGACATGCTTGGTGGAGAGCTGATTCGAAAGAGAGAGCAGAAATTG	960
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER		Db		280	LysIleLeuAlaHisAsnAsnPheValGlyArgLeuIleGlyLysGluGlyArgAsnLeu	299
; FILE REFERENCE: 210121.455C9		QY		961	AAGAAATTTGAACATGAACAGGACCAAGATTAACATCTCATCTTTGGAGGATTTGAGC	1020
; CURRENT APPLICATION NUMBER: US/09/606,421B		Db		300	LysLysIleGluGlnAspThrAspThrLysIleThrIleSerProLeuGlnGluLeuThr	319
; CURRENT FILING DATE: 2000-06-28		QY		1021	ATATACAAACCCGAAAGAACCATCATCTGTGAAGGGCACAGTGTGAGGCTGTGCCAGT	1080
; NUMBER OF SEQ ID NOS: 358		Db		320	LeuTyrAsnProGluArgThrIleThrValLysGlyAsnValGluThrCysAlaLysAla	339
; SOFTWARE: FastSeq for Windows Version 3.0		QY		1081	GAGATAGAGATTATGAGAGAGTCCGTGAGGCGCTTTGAAATATGATATGCTGGCTGTAA	1140
; SEQ ID NO 348		Db		340	GluGluGluIleMetLysLysIleArgGluSerTyrGluAsnAspIleAlaSerMetAsn	359
; LENGTH: 579		QY		1141	CAACAAGCCAACTGTATCCAGGGTGAACCTCAGCGCACTTGGCACTTTTCAACAGGA	1200
; TYPE: PR1		Db		360	LeuGlnAlaHisLeuIleProGlyLeuAsnLeuAsnAlaLeuGlyLeuPhe	376
; ORGANISM: Homo sapiens		QY		1201	CTCTCGTGTATCTCCACAGCAGGGCCCGCGAGCTCCCGCGCTGCCCTACCAC	1260
; US-09-606-421B-348		Db		377	-----ProThrSer-----GlyMetProProThr-----	386
Alignment Scores:		QY		1261	CCCTTCACTACCCACTCCGATCTCTCCAGCTGTACCCCATCACAGTTTGGCCCG	1320
Pred. No.:	4,02e-165	Length:	579	387	-----ProThrSer-----GlyMetProProThr-----Pro	396
Score:	1937.50	Matches:	579	1321	TTCCCGCATCATCTCTTATCCAGAGCAGGAGATTGTGAATCTCTTATCCCAACCCAG	1380
Percent Similarity:	79.19%	Conservative:	81	397	TyrProGlnPheGluGln---SerGluThrGluThrValHisLeuPheIleProAlaLeu	415
Best Local Similarity:	65.48%	Mismatches:	94	1381	GCTGTGGCGCCATCATCGGGAAGAGGGGCCACATCAACAGCTGGCGAGATTCCGC	1440
Query Match:	31.98%	Indels:	29	416	SerValGlyAlaIleIleGlyLysGlnGlyHisIleLysGlnLeuSerArgPheAla	435
DB:	4	Gaps:	8	1441	GGAGCTCTATCAAGATTGCCCTCGGAGGGCCGACGCTCAGCGAAAGATGGTTCATC	1500
US-09-270-437D-6 (1-3412) x US-09-606-421B-348 (1-579)		Db		436	GlyAlaSerIleLysIleAlaProAlaGluAlaProAspAlaLysValArgMetValIle	455
QY	73	ATGAAACAGCTTTACATCGGGAACCTGAGCCCGCCCGCTCACCGCGACGACCTCCGCGAG	132	1501	ATCACCGGGCCACCGGAAGCCAGTTCAGGCGCCAGGAGCGGATCTTTGGGAACTGAAA	1560
				456	IleThrGlyProProGluAlaGlnPheLysAlaGlnGlyArgIleTyrGlyLysIleLys	475
Db	1	MetAsnLysLeuTyrIleGlyAsnLeuSerGluAsnAlaAlaProSerAspLeuGluSer	20	1561	GAGGAAAATCTCTTTAAACCCAAAGAGAGTGAAGCTGAAGCGCATATCAGAGTGC	1620
				476	GluGluAsnPheValSerProLysGluGluValLysLeuGluAlaHisIleArgValPro	495
QY	133	CTCTTTGGGACAGGAGTCCCTCGCGGAGCAGGTCTGTGAAGTCCGGCTAGCC	192	1621	TCCTCCACAGCTGGCCGGGTGATTGGCAAGGTGGCAAGCCGTGAACGAACTGCAGAAC	1680
				496	SerPheAlaAlaGlyArgValIleGlyLysGlyLysThrValAsnGluLeuGlnAsn	515
Db	21	IlePheLysAspAlaLysIleProValSerGlyProPheLeuValLysThrGlyTyrAla	40	1681	TTAAACAGTGCAGAGTCTATCGTCTGTGACCAAAACCGACAGTCAAAATGAGGAAGTG	1740
				516	LeuSerSerAlaGluValValValProArgAspGlnThrProAspGluAsnAspGlnVal	535
QY	193	TTCTGTGACTACCCCGACAGAACTGGCCATCGCCCGCTCAGACCCGACGCTCCGGTAAA	252	1741	ATCGTCAGAAATATCGGCACTCTTTGTAGCCAGACTGCACAGCGCAGAGTACAGGAA	1800
				536	ValValLysIleThrGlyHisPheTyrAlaCysGlnValAlaGlnArgLysIleGlnGlu	555
Db	41	PheValAspCysProAspGluSerIleProValSerGlyProPheLeuValLysThrGlyTyrAla	60	1801	ATTGTACACAGGTGAGCAGCAGGAGCAGAAA	1833
QY	253	GTGGAATTGATGGGAAATCATGGAAGTTGATTACTCAGTCTCTAAAGAGCTAAGGAGC	312			
Db	61	IleGluLeuHisGlyLysProIleGluValGluHisSerValProLysArgGlnArgIle	80			
QY	313	AGGAAATTCAGATTGAAATATCCCTCTCCTCAGCTGGAGGAGTGTGGATGGACTT	372			
Db	81	ArgLysLeuGlnIleArgAsnIleProProHisLeuGlnIleProGluValLeuAspSerLeu	100			
QY	373	TTGGCTCAATATGGGACAGTGGAGAATGTGAACAGTCAACAGACAGACAGAAACCGCC	432			
Db	101	LeuValGlnTyrGlyValValGluSerCysGluGlnValAsnThrAspSerGluThrAla	120			
QY	433	GTTGTCACTCATATGCAACAGAGAGAGCAAAATAGCCATGGAGAGCTAAGC	492			
Db	121	ValValAsnValThrTyrSerSerLysAspGluAlaArgGlnAlaLeuAspLysLeuAsn	140			
QY	493	GGGCATCAGTTTGAGAACTACTCTTCAAGATTTCCTACATCCCGATGAGAGGTGAGC	552			
Db	141	GlyPheGlnLeuGluAsnPheThrLysValAlaTyrIleProAspGluThrAlaAla	160			
QY	553	TCCCTTTCGCCCTCAG-----CGAGCCAGCGT-----GGGGACCACTCTCCCGG	600			
Db	161	GlnGlnAsnProLeuGlnGlnProArgGlyArgArgGlyLeuGlnArgGlySerSer	180			
QY	601	GAGCAGGCCACGCCCTCGGGGCACTTCTCAGGCCAGACAGATGATTTCCCGCTCGCG	660			
Db	181	ArgGlnGly---SerProGlySerValSerLysGlnLysProCysAspLeuProLeuArg	199			
QY	661	ATCTGTCTCCCAACCCAGTTTGTGGTCCATCATCGGAAAGAGGGCTTGACATAAG	720			
Db	200	LeuLeuValProThrGlnPheValGlyAlaIleIleGlyLysGluGlyAlaThrIleArg	219			



QY 1621 TCTTCCACAGCTGGCCGGTGTATGGCAAGTGGCAAGCGTGAACGAACTGCAGAAC 1680  
 Db |||||  
 QY 496 SerPheAlaAlaGlyArgValIleGlyLysGlyLysThrValAsnGluLeuGlnAsn 515  
 Db |||||  
 QY 1681 TTAACCATGCAAGAGTCACTGCTGCTGACCAAGCCAGATCAAAATGAGGAAGTG 1740  
 Db |||||  
 QY 516 LeuSerAlaGluValValProArgAspGlnThrProAspGluAsnAspGlnVal 535  
 Db |||||  
 QY 1741 ATCCGTCAGATTATCGGCACCTCTTCTGCTAGCCAGACTGCACAGCGCAAGATCAGGAA 1800  
 Db |||||  
 QY 536 ValVallyIleThrGlyHisPheTyrAlaCysGlnValAlaGlnArgLysIleGlnGlu 555  
 Db |||||  
 QY 1801 ATTGTACACAGGTGAAGCAGCAGCAGCAAAA 1833  
 Db |||||  
 QY 556 IleLeuThrGlnValLysGlnHisGlnGlnGln 566

RESULT 6

US-09-480-884A-176  
 ; Sequence 176, Application US/09480884A  
 ; Patent No. 6482597  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Wang, Tongtong  
 ; APPLICANT: Fan, Liqun  
 ; APPLICANT: Hosken, Nancy A.  
 ; APPLICANT: Kalos, Michael D.  
 ; APPLICANT: Fanger, Gary R.  
 ; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY  
 ; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER  
 ; FILE REFERENCE: 210121.455C6  
 ; CURRENT APPLICATION NUMBER: US/09/480, 884A  
 ; NUMBER FILING DATE: 2001-08-27  
 ; NUMBER OF SEQ ID NOS: 330  
 ; SOFTWARE: FastSeq for Windows Version 3.0  
 ; SEQ ID NO 176  
 ; LENGTH: 579  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; US-09-480-884A-176

Alignment Scores:  
 Pred. No.: 1.69e-164 Length: 579  
 Score: 1950.50 Matches: 386  
 Percent Similarity: 79.02%  
 Best Local Similarity: 79.02%  
 Query Match: 31.87%  
 DB: 4  
 Gaps: 8

US-09-270-437D-6 (1-3412) x US-09-480-884A-176 (1-579)

QY 73 ATGAACAAGCTTTACATCGGGAACCTGAGCCCGCCGTCACCGCGACGACCTCCGCGAG 132  
 Db |||||  
 QY 1 MetAsnLysLeuTyrIleGlyAsnLeuSerGluAsnAlaAlaProSerAspLeuGluSer 20  
 QY 133 CTCTTTGGGACAGGAGTGGCCCTGGCGGACAGTCTGCTGAAGTCCGGCTACGCC 192  
 Db |||||  
 QY 21 IlePheLysAspAlaLysIleProValSerGlyProPheLeuValLysThrGlyTyrAla 40  
 QY 193 TTCGTGACTACCCGACCAAGACTGGCCCATCGGCGCATCGAGACCTCTCGGTAA 252  
 Db |||||  
 QY 41 PheValAspCysProAspGluSerTrpAlaLeuLysAlaIleGluAlaLeuSerGlyLys 60  
 QY 253 GTGGAATTGTCATGGGAAATCATGGAAGTGTACTTCAGTCTCTAAAGCTTAAGAGC 312  
 Db |||||  
 QY 61 IleGluLeuHisGlyLysProIleGluValGluHisSerValProLysArgGlnArgIle 80  
 QY 313 AGGAAATTCAGATTGAAACATCCCTCTCTACCTGAGTGGAGGTGTTGAGTGCAC 372  
 Db |||||  
 QY 81 ArgLysLeuGlnIleArgAsnIleProProHisLeuGlnTrpGluValLeuAspSerLeu 100  
 QY 373 TTGGCTCAATATGGACAGTGGAGATGTGGAACAAGTCAACACAGACAGAAACCGCC 432  
 Db |||||  
 QY 101 LeuValGlnTyrGlyValValGluSerCysGluGlnValAsnThrAspSerGluThrAla 120

QY 433 GTTGTCAACGTCACATATGCAACAAGAGAAGCAAAATAGCCATCGAGAGCTAAGC 492  
 Db |||||  
 QY 121 ValValAsnValThrTyrSerSerLysAspGlnAlaArgGlnAlaLeuAspLysLeuAsn 140  
 Db |||||  
 QY 493 GGCATCATGTTTGAGAACTACTCTTCAAGATTCTCTCATCCCGAGTGAAGAGTGAGC 552  
 Db |||||  
 QY 141 GlyPheGlnLeuGluAsnPheThrLeuLysValAlaTyrIleProAspGluMetAla 160  
 QY 553 TCCCTCTCGCCCTCAG-----CGAGCCAGCGT-----GGGAGCACCTCTTCCCGG 600  
 Db |||||  
 QY 161 GlnGlnAsnProLeuGlnGlnProArgGlyArgArgGlyLeuGlyGlnArgGlySer 180  
 QY 601 GAGCAAGCCACAGCCCTCGGGGCACTTCTCAGGCCACAGACATTGATTCCTCCGCTGGG 660  
 Db |||||  
 QY 181 ArgGlnGly--SerProGlySerValSerLysGlnLysProCysAspLeuProLeuArg 199  
 QY 661 ATCCTGCTCCCGCCAGTTTGTGGTCCCATCATCGGAAAGGAGGGCTTGACCATAAAG 720  
 Db |||||  
 QY 200 LeuLeuValProThrGlnPheValGlyAlaIleIleGlyLysGluGlyAlaThrIleArg 219  
 QY 721 AACATCACTAAGCAGACCCAGTCCCGGTAGATATCCATAGAAAGAGAACTCTGGAGCT 780  
 Db |||||  
 QY 220 AsnIleThrLysGlnThrGlnSerLysIleAspValHisArgLysGluAsnAlaGlyAla 239  
 QY 781 GCAGAGAAGCTGTCACTCCATCCATCCACCCAGGGGACTTCTGAAGCATGCGCGATG 840  
 Db |||||  
 QY 240 AlaGluLysSerIleThrIleLeuSerThrProGluGlyThrSerAlaAlaCysLysSer 259  
 QY 841 ATTCTTGAATCATGCAAGAGGAGGACAGTGAACCAACTAGCCGAGAGATTCTCTGTG 900  
 Db |||||  
 QY 260 IleLeuGluIleMetHisLysGluAlaGlnAspIleLysPheThrGluGluIleProLeu 279  
 QY 901 AAAATCTTGGCACACAATGCTGTTGGAAGACTGATGCAAAAGAGGAGCAAAATTG 960  
 Db |||||  
 QY 280 LysIleLeuAlaHisAsnAsnAsnPheValGlyArgLeuIleGlyLysGluGlyArgAsnLeu 299  
 QY 961 AAGAAATTGAACATGAACAGGAGCAAGATAACAATCTCATCTTTSCAGGATTTGAGC 1020  
 Db |||||  
 QY 300 LysLysIleGluGlnAspThrAspThrLysIleThrIleSerProLeuGlnLeuThr 319  
 QY 1021 ATATACAACCCGGAAGAACCATCACTGTGAAGGACAGTGTGAGGCTGTGCCAGTGCT 1080  
 Db |||||  
 QY 320 LeuTyrAsnProGluArgThrIleThrValLysGlyAsnValGluThrCysAlaLysAla 339  
 QY 1081 GAGATAGAGATTATGAAGAAGTCCGTGAGCGCTTTGAAATGATATGCTGGCTGTAAAC 1140  
 Db |||||  
 QY 340 GluGluGluIleMetLysLysIleArgGluSerTyrGluAsnAspIleAlaSerMetAsn 359  
 QY 1141 CAACAAGCCAACTGTATCCAGGGTTGAACCTCAGCGCACTTGGCATCTTTTCAACAGGA 1200  
 Db |||||  
 QY 360 LeuGlnAlaHisLeuIleProGlyLeuAsnLeuAsnAlaLeuGlyLeuPhe----- 376  
 QY 1201 CTGTCGCTGTACTCTCCACAGCAGGCGCCCGGAGCTCCCGCGCTGCCCTACCAC 1260  
 Db |||||  
 QY 377 -----ProThrSer-----GlyMetProProThr----- 386  
 QY 1261 CCCTTCACTACCCACTCCGATACTTCTCCAGCTGTACCCCATCACCAGTTTGGCCCG 1320  
 Db |||||  
 QY 387 -----SerGlyProProSerAlaMetThrPro-----Pro 396  
 QY 1321 TTCCCGCATCATCTCTTATCCAGAGCAGGAGATTGTGAATCTTCTTATCCCAACCCAG 1380  
 Db |||||  
 QY 397 TyrProGlnPheGluGln---SerGluThrGluThrValHisGlnPheIleProAlaLeu 415  
 QY 1381 GCTGTGGCGCCCATCATCGGGAAGAGGGGCGACACATCAACAGCTGCGGAGATTGCGC 1440  
 Db |||||  
 QY 416 SerValGlyAlaIleIleGlyLysGlnGlnHisIleLysGlnLeuSerArgPheAla 435  
 QY 1441 GGAGCTCTTATCAAGATTGCCCCCTCGGGAAGCCAGAGCTCAGCGAAGAGGATGTCATC 1500  
 Db |||||  
 QY 436 GlyAlaSerIleLysIleAlaProAlaGluAlaProAspAlaLysValArgMetValIle 455  
 QY 1501 ATCAGCGGCGCACCGGAAGCCCGAGTTCAGGCGCCGAGCGGATCTTTGGGAAACTGAAA 1560

456	IleThrGlyProGluAlaGlnPheLysAlaGlnGlyArgIleThrGlyLysIleLys	475
1561	GAGGAAACCTCTTTAAACCCCAAGAGAGTGAAGCTGCAAGCGCATATCAGAGTGCCC	1620
476	GluGluAsnPheValSerProLysGluGluValLysLeuAlaHisIleArgValPro	495
1621	TCTTCCACAGCTGGCGGGTGATGTGCAAGGTGGCAAGCCGTGAACGAATCCAGAAC	1680
496	SerPheAlaAlaGlyArgValIleGlyLysGlyLysThrValAsnGluLeuGlnAsn	515
1681	TTAAACAGTGCAGAATCATCTGCTCGTGACCAACGCCAGATGAAATGAGGAAGTG	1740
516	LeuSerSerAlaGluValValProArgAspGlnThrProAspGluAsnAspGlnVal	535
1741	ATCGTCCAGATTATTCGGGCCTCTTTGTCTAGCCAGACTCCACAGCCCAAGATCAGGAA	1800
536	ValValLysIleThrGlyHisPheThrAlaCysGlnValAlaGlnArgLysIleGlnGlu	555
1801	ATTGTACAACAGGTGAAGCAGGAGCAGAA	1833
556	IleLeuThrGlnValLysGlnHisGlnGlnGln	566

## RESULT 7

```

US-09-542-615A-176
; Sequence 176, Application US/09542615A
; Patent No. 6518256
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy A.
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C8
; CURRENT APPLICATION NUMBER: US/09/542,615A
; CURRENT FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 350
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 176
; LENGTH: 579
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-542-615A-176

```

Alignment Scores:	
Pred. No.:	1.69e-164
Scores:	1950.50
Percent Similarity:	75.01%
Best Local Similarity:	69.32%
Query Match:	31.87%
DR:	4
Length:	579
Matches:	386
Conservative:	81
Mismatches:	95
Indels:	29
Gaps:	8

US-09-270-437D-6 (1-3412) x US-09-542-615A-176 (1-579)

QY	73	ATGAACAAAGCTTTTACATCGGGAACCTTGAGCCGCCGCCGTCA	CGCGCGAGACGACCTCGGCAG	132
Db	1	MetAsnIysLeuTyrlleGlyAsnLeuSerGluAsnAlaAlaProSerAsp	LeuGluSer	20
QY	133	CTCTTTGGGACAGAAAGCTCCCTCGCGGAGACAGGTCCTGCTG	AGCTTCGGCTACGCC	192
Db	21	IlePheIysAspAlaIysileproValSerGlyProPheLeuValIysThr	GlyTyrAla	40
QY	193	TTCTGTGAAGTACCCCGACCAAGTGGCCCATTCGCGCCATCGAG	ACACCTCTCGGGTAAA	252
Db	41	PheValAspCysProAspGluSerTrpAlaLeuIysAlaIleGluAla	LeuSerGlyIys	60
QY	253	GTGGAAATTCATGGGAAATCATGGAAGTTGATTACTCAGTCTCT	TAATAAGCTAAGGAC	312
Db	61	IleGluLeuHisGlyIysProIleGluValGluHisSerValProIysArg	GlnIleGln	80

Db 416 SerValGlyAlaIleGlyGlnGlnHisIleLysGlnLeuSerArgPheAla 435  
QY 1441 GGAGCTCTATCAAGATTGCCCTCGGAAGCCAGAGCTCAGCGAAGGATGTCATC 1500  
Db 436 GlyAlaSerIleLysIleAlaProAlaGluAlaProAspAlaLysValArgMetValIle 455  
QY 1501 ATCCCGGGCCACCGGAGCCAGTTCAGGCCCGGAGGACGGATTTGGGAACTGAAA 1560  
Db 456 IleThrGlyProGluGlnGlnPheLysAlaGlnGlyArgIleTyrGlyLysIleLys 475  
QY 1561 GAGGAAACTTCTTTAAACCCCAAGAAAGTGAAGCTGGAAGCCATATCAGAGTGC 1620  
Db 476 GluGluAsnPheValSerProLysGluGluValLysLeuGluAlaHisIleArgValPro 495  
QY 1621 TCTTCCACAGCTGGCGGGTGATTTGGGAAAGTGGCAAGCCGTCAGAACTGCAAG 1680  
Db 496 SerPheAlaAlaGlyArgValIleGlyLysGlyGlyLysThrValAsnGluLeuGlnAsn 515  
QY 1681 TTAACCACTGAGAGTCTATCGTCTGTCACCAAGCCAGATGAAATGAGGAAGTG 1740  
Db 516 LeuSerSerAlaGluValValProArgAspGlnThrProAspGluAsnAspGlnVal 535  
QY 1741 ATCGTCAGATTATGGGCACTTCTTCTAGCCAGACTGCACAGCGCAAGATCAGGAA 1800  
Db 536 ValValLysIleThrGlyHisPheTyrAlaCysGlnValAlaGlnArgLysIleGlnGlu 555  
QY 1801 ATTGTACAAAGTGTAGCAGCAGAGCAGAGAA 1833  
Db 556 IleLeuThrGlnValLysGlnHisGlnGlnGln 566

## RESULT 8

US-09-606-421B-176  
; Sequence 176, Application US/09606421B  
; Patent No. 6531315  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Tongtong  
; APPLICANT: Fan, Liqun  
; APPLICANT: Kalos, Michael D.  
; APPLICANT: Bangur, Chaitanya S.  
; APPLICANT: Hosken, Nancy  
; APPLICANT: Fanger, Gary R.  
; APPLICANT: Li, Samuel X.  
; APPLICANT: Wang, Aijun  
; APPLICANT: Skeiky, Yasir A.W.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER  
; FILE REFERENCE: 210121.455C9  
; CURRENT APPLICATION NUMBER: US/09/606,421B  
; CURRENT FILING DATE: 2000-06-28  
; NUMBER OF SEQ ID NOS: 358  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 176  
; LENGTH: 579  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-606-421B-176

Alignment Scores:  
Pred. No.: 1.69e-164 Length: 579  
Score: 1950.50 Matches: 386  
Percent Similarity: 79.02% Conservative: 81  
Best Local Similarity: 65.31% Mismatches: 95  
Query Match: 31.87% Indels: 29  
DB: 4 Gaps: 8

US-09-270-437D-6 (1-3412) x US-09-606-421B-176 (1-579)

QY 73 ATGACAGCTTTACATCGGAACCTGAGCCCGCCCTCACCGCCGACGACCTCCGCGAG 132  
Db 1 MetAsnLysLeuTyrIleGlyAsnLeuSerGluAsnAlaAlaProSerAspLeuGluSer 20  
QY 133 CTCTTTGGGACAGGAAGCTGCCCTCGGGGACAGGTCCTGCTGAAGTCCGGCTAGGCC 192

Db 21 IlePheLysAspAlaLysIleProValSerGlyProPheLeuValLysThrGlyTyrAla 40  
QY 193 TTCTGGGACTACCCGACGACAGAACTGGGCCATCCGCGCCATCGACACCTCTCTCGGGTAAA 252  
Db 41 PheValAspCysProAspGluSerTrpAlaLeuLysAlaIleGluAlaLeuSerGlyLys 60  
QY 253 GTGGAATTGCATGGGAAATCATGGAAGTTGATTACTCAGTCTCTTAAAGAGCTAAGGAGC 312  
Db 61 IleGluLeuHisGlyLysProIleGluValGluHisSerValProLysArgGlnArgIle 80  
QY 313 AGGAAATTCAGATTCGAAACATCCTCTCCTCAGCTGAGTGGAGGTGTGATGAGCTT 372  
Db 81 ArgLysLeuGlnIleArgAsnIleProProHisLeuGlnTrpGluValLeuAspSerLeu 100  
QY 373 TTGGCTCAATATGGACAGTGCAGAACTGTGAAACAAGTCAACACAGACACAGAACCCGCC 432  
Db 101 LeuValGlnTyrGlyValValGluSerCysGluGlnValAsnThrAspSerGluThrAla 120  
QY 433 GTTGTCAACGTTCACATATGCAACAGAGAAAGCAAAATAGCCATCGAGAAAGCTAAGC 492  
Db 121 ValValAsnValThrTyrSerSerLysAspGlnAlaArgGlnAlaLeuAspLysLeuAsn 140  
QY 493 GGGCATCAGTTTGAGAACTACTCCTTCAAGATTTCTTACATCCCGAGTGAAGAGTGAAG 552  
Db 141 GlyPheGlnLeuGluAsnPheThrLeuLysValAlaTyrIleProAspGluMetAlaAla 160  
QY 553 TCCCTTCGCCCCCTCAG-----CGAGCCAGAGCT-----GGGACCACCTCTCCCGG 600  
Db 161 GlnGlnAsnProLeuGlnGlnProArgLysArgGlyLeuGlyGlnArgGlySerSer 180  
QY 601 GAGCAAGCCACGCCCCCTCGGGGCACTTCTCAGCCACAGACAGATTTGCCCTCGCG 660  
Db 181 ArgGlnGly--SerProGlySerValSerLysGlnLysProCysAspLeuProLeuArg 199  
QY 661 ATCCTGCTCCCGCCCGGATTTGTTGGTCCCATCATCGGAAGAGGGCTTGACCATAAAG 720  
Db 200 LeuLeuValProThrGlnPheValGlyAlaIleGlyLysGlyLysGlyAlaThrIleArg 219  
QY 721 AACATCACTAAGACAGACCCAGTCCCGGTAGATATCCATAGAAAGAGAGACTCTGGAGCT 780  
Db 220 AsnIleThrLysGlnThrGlnSerLysIleAspValHisArgLysGluAsnAlaGlyAla 239  
QY 781 GCAGAGAGGCTGTCAACATCCATCCACCCCGGAGGAGCTTCTGAAGATTCGCGCATG 840  
Db 240 AlaGluLysSerIleThrIleLeuSerThrProGluGlyThrSerAlaAlaCysLysSer 259  
QY 841 ATTCTTGAATCATGCAAGAGGAGGAGGAGTGAACCAACTAGCCGAGAGATTCTCTGTG 900  
Db 260 IleLeuGluIleMetHisLysGluAlaGlnAspIleLysPheThrGluGluIleProLeu 279  
QY 901 AAAATCTTGGCACACAAATGCTTGGTTGGAAGACTGATTGGAAGAGGAGGAGCAAAATTG 960  
Db 280 LysIleLeuAlaHisAsnAsnPheValGlyArgLeuIleGlyLysGlyLysArgGlnLeu 299  
QY 961 AAGAAATTCGAACATGAACAGAGGACCAAGATAACATCTCATCTTTGCGAGATTTCAGC 1020  
Db 300 LysLysIleGluGlnAspThrAspThrLysIleThrLysSerProLeuGlnGluLeuThr 319  
QY 1021 ATATACACCGGAAAGAACCATCATCTGTGAAGGGCAGACTTGAAGGCTGTGCGAGTCT 1080  
Db 320 LeuTyrAsnProGluArgThrIleThrValLysGlyAsnValGluThrCysAlaLysAla 339  
QY 1081 GAGATAGAGATTATGAAGAGCTGCTGAGGCGCTTTGAAATCATATGCTGGCTGTGTAAAC 1140  
Db 340 GluGluGluIleMetLysLysIleArgGluSerTyrGluAsnAspIleAlaSerMetAsn 359  
QY 1141 CAAACAAGCAATCTGATCCCGGTTGAACCTCAGCGCACTTGGCATCTTTTCAACAGGA 1200  
Db 360 LeuGlnAlaHisLeuIleProGlyLeuAsnLeuAsnAlaLeuGlyLeuPhe----- 376  
QY 1201 CTCTCCGTGCTATCTCACACAGCAGGCGGCCCGGAGCTGCCCGCTGCCCTTACCAC 1260



Db 377 -----ProProThrSer-----GlyMetProProProThr----- 386  
QY 1261 CCCTTCACTACCACTCGGATACCTTCACGCGTGTACCCCATCACAGTTTGGCGG 1320  
Db 387 -----SerGlyProProSerAlaMetThrPro-----Pro 396  
QY 1321 TTCGCGCATCATCTTATCCAGACGAGGAGATTGTAATCTCTTCATCCACCCAG 1380  
Db 397 TyrProGlnPheGln-----SerGluThrGluThrValHisGlnPheProAlaLeu 415  
QY 1381 GCTGTGGCGCCATCATCGGAAGAGGGGACACATCAACAGCTGCGAGATTCGCC 1440  
Db 416 SerValGlyAlaIleleelysGlnGlyGlnHisIleLysGlnLeuSerArgPheAla 435  
QY 1441 GGAGCTCTATCAAGATTGCCCTGCGGAAGGCCAGACCTCAGCGAAGGATGTCATC 1500  
Db 436 GlyAlaSerIleLeyleAlaProAlaGluAlaProAspAlaLysValArgMetValIle 455  
QY 1501 ATCACCAGGCGCACCGGAAGCCAGTTCAGGCGCCAGGACGATCTTTGGGAAACTGAAA 1560  
Db 456 IleThrGlyProProGluAlaGlnPheLysAlaGlnGlyArgIleTyrGlyLysIleLys 475  
QY 1561 GAGGAAACTCTTTAACCCCAAGAGAGTGAAGCTGGAAGCGCATATCAGATGCC 1620  
Db 476 GluGluAsnPheValSerProLysGluGluValLysLeuGluAlaHisIleArgValPro 495  
QY 1621 TCTTCCAGCTGGCGGGTGTATTGGCAAGCTGGCAGACCGTGAACGAACTGCAGAAC 1680  
Db 496 SerPheAlaAlaGlyArgValIleGlyLysGlyLysValAsnGluLeuGlnAsn 515  
QY 1681 TTAACCCAGTGCAAGTTCATCGTCTGCTGACCAAAACCCAGATGAAATGAGGAAGTG 1740  
Db 516 LeuSerSerAlaGluValValProArgAspGlnThrProAspGluAsnAspGlnVal 535  
QY 1741 ATCGTCAGATATTCGGCACTCTTTGTAGCCAGCTGCACAGCGCAAGATCAGGGAA 1800  
Db 536 ValValLysIleThrGlyHisPheTyrAlaCysGlnValAlaGlnArgLysIleGlnGlu 555  
QY 1801 ATTGTACACAGCTGAAGCAGCAGGAGCAGAAA 1833  
Db 556 IleLeuThrGlnValLysGlnHisGlnGlnGln 566

## RESULT 9

US-08-021-608D-2  
Sequence 2, Application US/08021608D

Patent No. 5580760

## GENERAL INFORMATION:

APPLICANT: LEVENS, DAVID L., DUNCAN,  
APPLICANT: ROBERT C., AND AVIGAN, MARK I.  
TITLE OF INVENTION: NOVEL FUSE BINDING  
TITLE OF INVENTION: PROTEIN AND CDNA THEREFOR  
NUMBER OF SEQUENCES: 24

## CORRESPONDENCE ADDRESS:

ADDRESSEE: MORGAN & FINNEGAN  
STREET: 345 PARK AVENUE  
CITY: NEW YORK  
STATE: NEW YORK  
COUNTRY: USA  
ZIP: 10154

## COMPUTER READABLE FORM:

MEDIUM TYPE: FLOPPY DISK  
COMPUTER: IBM PC COMPATIBLE  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WORDPERFECT 5.1

## CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/021,608D  
FILING DATE: 22-FEB-1993  
CLASSIFICATION: 435

## PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE: 435

## CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: WILLIAM S. FEILER  
REGISTRATION NUMBER: 26,728  
REFERENCE/DOCKET NUMBER: 2026-4063  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 758-4800  
TELEFAX: (212) 751-6849  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 644  
TYPE: Amino Acid  
STRANDEDNESS: Single  
TOPOLOGY: Unknown  
MOLECULE TYPE: Peptide/Protein  
HYPOTHETICAL: No  
ORIGINAL SOURCE:  
ORGANISM: Human  
CELL LINE: HL60  
FEATURE:  
OTHER INFORMATION: Amino Acid 149 (Xaa) is Met or Ile  
US-08-021-608D-2

## Alignment Scores:

Pred. No.: 6,42e-13 Length: 644  
Score: 245.00 Matches: 145  
Percent Similarity: 35.64% Conservative: 82  
Best Local Similarity: 22.76% Mismatches: 241  
Query Match: 4.00% Indels: 169  
DB: 1 Gaps: 27

US-09-270-437D-6 (1-3412) x US-08-021-608D-2 (1-644)

QY 336 CCTCTCTCACCTCGAGTGGAGGTCTTGATGGACTTTTGGCTCAATATGGACAGTGA 395  
Db 10 ProSerSerGlySerAlaGlyGlyGly-----GlyGlyGly 22  
QY 396 GAATGTGAACAAGTCAACACACACAGACAAACCCCGTGTCAACGTGCATATGCAAC 455  
Db 23 GlyGlyGly-GlyValAsnAspAlaPheLysAspAlaLeuGlnArg-----AlaAr 39  
QY 456 AAGAGAAGACAAATAAGCCATCGAGAGCTAGCGGGCATCAGTTGTGAGAACTACTC 515  
Db 39 gGlnIleAlaAlaLysIleGlyGlyAspAlaGlyThrSerLeuAsnSerAsnAspTyrGl 59  
QY 516 CTTCAAGATTCTCATCTCCGATGAAGAGGTGAGCTCCCTTCGCCCTCAGCGAGC 575  
Db 59 yTyrGlyGlyGlnLysArgProLeuGluAspGlyAspGlnProAspAlaLysLysVal 79  
QY 576 CCAGCTGGGACCACTCTTCCCGGAGCAAGCCACGCCCTCTGGGGGCACTTCTCAGGC 635  
Db 79 aProGlnAsnAspSerPheGlyThrGlnLeu-----ProProMetHisGlnGlnGlnSe 97  
QY 636 CAGACAGATTGATTTCCCGCTCGGATCTGTGTCGCCACCCAGCTTGTGTGGTGCATCAT 695  
Db 97 rArgSerVal---MetThrGluGluTyrLysValProAspGlyMetValGlyPheIle 116  
QY 696 CGGAAAGGAGGCTTCACCATAAAGACATCTACTAAGCAGACCCAGTCCCGGTAGATAT 755  
Db 116 eGlyArgGlyGlyGluGlnIleSerArgIleGlnGlnGluSerGlyCysLysIleGlnI 136  
QY 756 CCATGAAAGAGAACTCTGGAGCTGCAGAGAAGCCGTGCACCATCCATCCGCCACCCAGA 815  
Db 136 e---AlaProAspSerGlyGlyLeuProGluArgSerCys\*\*\*LeuThrGlyThrProGl 155  
QY 816 GGGGATCTCTGAAGCATGCCGATGATCTTCAATCATGACGAAA----- 861  
Db 155 uSerValGlnSerAlaLysArgLeuLeuAspGlnIleValGluLysGlyArgProAlaPr 175  
QY 862 -----GAGGCGAGTACAGACCAACTAGCCGAAAGATTCCTCTGAAATCTTGGC 911  
Db 175 oGlyPheHisHisGlyAspGlyProGlyAsnAlaValGln-----GluIleMetIl 192  
QY 912 ACAATGCTTGGTGGAGAGTATTGGAAAAGAGGACAGAAAATTTGAGAGAAATGGA 971



[illegible]

Db	475	roGly- ProGly-	----	ThrPro-	-----MetGlyPro	483
QY	1938	AACGCAGCCACCCAGATCGGGAGCAAAACCAAGACCATCTGAGGAATGAGAAGTCTCGGG	1997			
Db	484	TyrAsnProAlaProTyrAsnProGlyPro-	----		----	493
QY	1998	AGCGGGCCAGGACTCTGCGGAGG-CCCTGAGAACCCAGGGCCGAGGAGGGCGGG	2054			
Db	494	-----ProGlyProAlaProHisGlyProProAlaProTyrAlaProGlnGlyTrpGly	511			
QY	2055	GAAGGTAGCCAGGTTGCCAGAA-----CCACCGAGCCCCGGC	2093			
Db	512	AsnAlaTyrProHisTrpGlnGlnAlaProProAspProAla	526			
RESULT 10						
US-08-726-160-2						
; Sequence 2, Application US/08726160						
; Patent No. 5734016						
; GENERAL INFORMATION:						
; APPLICANT: LEVENS, DAVID L., DUNCAN,						
; APPLICANT: ROBERT C., AND AVIGAN, MARK I.						
; TITLE OF INVENTION: NOVEL FUSE BINDING						
; TITLE OF INVENTION: PROTEIN AND CDNA THEREFOR						
; NUMBER OF SEQUENCES: 24						
; CORRESPONDENCE ADDRESS:						
; ADDRESSEE: MORGAN & FINNEGAN						
; STREET: 345 PARK AVENUE						
; CITY: NEW YORK						
; STATE: NEW YORK						
; COUNTRY: USA						
; ZIP: 10154						
; COMPUTER READABLE FORM:						
; MEDIUM TYPE: FLOPPY DISK						
; COMPUTER: IBM PC COMPATIBLE						
; OPERATING SYSTEM: PC-DOS/MS-DOS						
; SOFTWARE: WORDPERFECT 5.1						
; CURRENT APPLICATION DATA:						
; APPLICATION NUMBER: US/08/726,160						
; FILING DATE: 04-OCT-1996						
; CLASSIFICATION: 435						
; PRIOR APPLICATION DATA:						
; APPLICATION NUMBER: 08/021,608						
; FILING DATE: 22-FEB-1993						
; CLASSIFICATION: 435						
; ATTORNEY/AGENT INFORMATION:						
; NAME: WILLIAM S. FEILER						
; REGISTRATION NUMBER: 26,728						
; REFERENCE/DOCKET NUMBER: 2026-4063US1						
; TELECOMMUNICATION INFORMATION:						
; TELEPHONE: (212) 758-4800						
; TELEFAX: (212) 751-6849						
; INFORMATION FOR SEQ ID NO: 2:						
; SEQUENCE CHARACTERISTICS:						
; LENGTH: 644						
; TYPE: Amino Acid						
; STRANDEDNESS: Single						
; TOPOLOGY: Unknown						
; MOLECULE TYPE: Peptide/Protein						
; HYPOTHETICAL: No						
; ORIGINAL SOURCE:						
; ORGANISM: Human						
; CELL LINE: HL60						
; FEATURE:						
; OTHER INFORMATION:						
; OTHER INFORMATION: Amino Acid 149 (Xaa) is Met or Ile						
US-08-726-160-2						
Alignment Scores:						
Pred. No.: 6 42e-13 Length: 644						
Score: 245.00 Matches: 145						
Percent Similarity: 35.64% Conservative: 82						
Best Local Similarity: 22.76% Mismatches: 241						
Query Match: 4.00% Indels: 169						









US-09-270-437D-6 (1-3412) x US-08-726-160-10 (1-643)

QY 336 CCCTCCTACCTGCGAGTGGAGGTGTGGATGGACTTTTGGCTCAATATGGGACAGTGA 395  
 Db ||||| : : : ||||| ||  
 10 ProSerGlySerAlaGlyGlyGly-----GlyGlyGly 22  
 QY 396 GAATGTGAACAAGTCAACACAGACACACAGACACACCGCGTGTCAACGTCACATATGCAAC 455  
 Db ||||| : : : ||||| ||  
 23 GlyGlyGly-GlyValAsnAspAlaPheLysAspAlaLeuGlnArg-----AlaAr 39  
 QY 456 AAGAGAAGAGCAAAATAGCCATGGAGAGCTAAGCGGCGCATGTTGTGAGAACTATCTC 515  
 : : : ||||| : : : ||||| ||  
 Db 39 gGlnleAlaAlaLysIleGlyGlyAspAlaGlyThrSerLeuAsnSerAsnAspTyrGl 59  
 QY 516 CTTCAAGATTTCTACATCCCGATGACAGAGTGGAGTCCCTTCGCCCCCTCAGCGAGC 575  
 : : : ||||| : : : ||||| ||  
 Db 59 yTyrGlyGlyGlnLysArgProLeuGluAspGlyAspGlnProAspAlaLysLysValAl 79  
 QY 576 CCAGCGTGGGACCACTCTTCCCGGAGCAAGGCCACCGCCCTGGGGGCACCTTCTCAGGC 635  
 : : : ||||| : : : ||||| ||  
 Db 79 aProGlnAsnAspSerPheGlyThrGln-----LeuProMetHisGlnGlnGl 96  
 QY 636 CAGACAGATGATTTCCCGCTCGGATCCTGTGTCGCCACCCAGTGTGTGTGGTCCATCAT 695  
 ||||| : : : ||||| ||  
 Db 96 nArgSerVal---MetThrGluLutyrLysValProAspGlyMetValGlyPheIleIl 115  
 QY 696 CGGAAGGAGGGCTTGACCATAAAGACATCACTAAGCAGACCCAGTCCCGGTAGATAT 755  
 ||||| : : : ||||| : : : ||||| ||  
 Db 115 eGlyArgGlyGlyGluGlnIleSerArgIleGlnGlnLysGlyCysLysIleGlnIl 135  
 QY 756 CCATGAAAGAGAACTCTGAGCTGCAGAGAGCTGTCAACCATCCATGCCACCCACAGA 815  
 : : : ||||| : : : ||||| ||  
 Db 135 e---AlaProAspSerGlyLeuProGluArgSerCys\*\*LeuThrGlyThrProGl 154  
 QY 816 GGGACTCTGAGCATGCCGATGATTTCTTGAATCATGCGAGAA 861  
 : : : ||||| : : : ||||| ||  
 Db 154 uSerValGlnSerAlaLysArgLeuLeuAspGlnIleValGluLysGlyArgProAlaPr 174  
 QY 862 -----GAGGCAGATGAGACCAAACTAGCGAGAGATCTCTCTGAAATCTTGGC 911  
 ||||| : : : ||||| : : : ||||| ||  
 Db 174 oGlyPheHisGlyAspGlyProGlyAsnAlaValGln-----GluIleMetIl 191  
 QY 912 ACACAATGGCTGTGTTGAAGACTGATTGGAAGAAGGCAGAAATTTGAAGAAATTTGA 971  
 ||||| : : : ||||| : : : ||||| ||  
 Db 191 eProAlaSerLysAlaGlyLeuValIleGlyLysGlyGlyGluThrIleLysGlnLeuGl 211  
 QY 972 ACATGAACAGAGCACAGATAACTATCTCTTTGAGGATTTGAGCATATACACCC 1031  
 : : : ||||| : : : ||||| ||  
 Db 211 nGluArgAlaGlyValLysMetValMet-----IleGlnAsp-----GlyPr 225  
 QY 1032 GGAAGAACCATCACTGTGAAGGCACAGTGTGAGGCTGTGCGAGTGTGAGATAGAGAT 1091  
 : : : ||||| : : : ||||| ||  
 Db 225 oGlnAsnThr-----GlyAlaAsp----- 231  
 QY 1092 TATGAAGAGTGCCTGAGGCTTTGAAATGATATGCTGGTGTGTTAAACCAACAGCCAA 1151  
 ||||| : : : ||||| ||  
 Db 232 ---LysProLeuArg-----IleThrGlyAspProTyrLysValGlnGlnAlaLysGl 248  
 QY 1152 TCTGATCCCGGTGAACCTCAGGCACCTTGGCATCTTTTCAACAGGACTGTCCGTGCT 1211  
 : : : ||||| : : : ||||| ||  
 Db 248 uMetVal-----LeuGluLeuIl 254  
 QY 1212 ATCTCCACAGAGGCGCGGAGCTCCCGCGTGTGCGCTTACCACCCCTTCACTAC 1271  
 : : : ||||| : : : ||||| ||  
 Db 254 eArgAspGlnGlyPheArgGluVal----- 263  
 QY 1272 CCACCTCCGATACTTCTCAGCGCTGTACCCCATACACAGTGTGGCCCTTCCCGCATCA 1331  
 : : : ||||| : : : ||||| ||  
 Db 264 -----ArgAsnGluTyrGly----- 268  
 QY 1332 TCACTTTATCCAGACAGAGATTTGTAATCTTTCATCCCAACCCAGGCTGTGGCGC 1391  
 : : : ||||| : : : ||||| ||  
 Db 269 -SerArgIleGlyGlyAsnGluGlyIleAspValProIleProArgPheAlaValGlyIl 288

QY 1392 CATCATCGGGAAGGGGCACACATCAAAACAGCTGGCAGATTTCGCGGAGCCTCTAT 1451  
 : : : ||||| : : : ||||| : : : ||||| ||  
 Db 288 eValIleGlyArgAsnGlyGluMetIleLysIleGlnAsnAspAlaGlyValArgIl 308  
 QY 1452 CAAGATTCCTCCCTCGGAAGGCCACAGCTCAGCAAAAGGATGTCATCATCACCGGCGC 1511  
 : : : ||||| : : : ||||| : : : ||||| ||  
 Db 308 eGlnPheLysProAspAspGly---ThrThrProGluArgIleAlaGlnIleThrGlyPr 327  
 QY 1512 ACCGAA---GCCAGTTCAAGGCCAG----- 1536  
 : : : ||||| : : : ||||| ||  
 Db 327 oProAspArgCysGlnHisAlaAlaGluIleThrAspLeuLeuArgSerValGlnAl 347  
 QY 1537 -----GGACGGATCTTTGGAAA----- 1554  
 : : : ||||| : : : ||||| ||  
 Db 347 aGlyAsnProGlyGlyProGlyProGlyGlyArgGlyArgGlyGlnGlyAsnTr 367  
 QY 1555 -----CTGAAAGAGAAATCTTTTAAACCCCAAGAGAGT 1592  
 : : : ||||| : : : ||||| ||  
 Db 367 pAsnMetGlyProProGlyGlyLeuGlnGluPheAsnPheIlle----- 381  
 QY 1593 GAAGCTGAAGCGCATATCAGAGTGCCTCTTCCACAGCTGCCCGGTGATTGGCAAGG 1652  
 : : : ||||| : : : ||||| : : : ||||| ||  
 Db 382 -----ValProThrGlyLysThrGlyLeuIleIleGlyLysGl 394  
 QY 1653 TGGCAAGCCGTGAACGAACCTGCAAGCTTAACCAAGTCAGAAAGTCATCGTCCCTCGTGA 1712  
 : : : ||||| : : : ||||| : : : ||||| ||  
 Db 394 yGlyGluThrIleLysSerIleSerGlnGlnSerGlyAlaArgIleGluLeuGlnArgAs 414  
 QY 1713 CCAAGCCGATGAAATAGGAAGTG---ATCCTCAGAAATATTCGGGCAC----- 1761  
 : : : ||||| : : : ||||| : : : ||||| ||  
 Db 414 nProProAsnAlaAspProAsnMetLysLeuPheThrIleArgGlyThrProGlnGl 434  
 QY 1762 ----TTCCTTGTACCCAGACTGTCAGCGGAGTGCCTTCACAGCGCAGCAAGTGAGGCTCCCA 1877  
 : : : ||||| : : : ||||| : : : ||||| ||  
 Db 434 nIleAspTyrAlaArgGlnLeuIleGluLysIleGlyGlyProValAsnProLeuGl 454  
 QY 1818 GCAGCAGGAGCAGAAATACCTTCAGGAGTGCCTTCACAGCGCAGCAAGTGAGGCTCCCA 1877  
 : : : ||||| : : : ||||| : : : ||||| ||  
 Db 454 yProProValProHisGlyProHisGlyVal-ProGlyProHisGlyProGlyProp 474  
 QY 1878 CAGGCACGAGCAAAACAAGGATGAATGTAGCCCTTCCACACCTCAGACAGATGAGACCA 1937  
 : : : ||||| : : : ||||| : : : ||||| ||  
 Db 474 roGly-ProGly-----ThrPro-----MetGlyPro 482  
 QY 1938 AACGAGCAGCCAGCTCGGAGCAACCAAGAACCATCTGAGGAATGAGAAGTCTGCGG 1997  
 : : : ||||| : : : ||||| : : : ||||| ||  
 Db 483 TyrAsnProAlaProTyrAsnProGlyPro----- 492  
 QY 1998 AGGCGCGCAGGACTCTGCCGAGG---CCCTGAGAACCCCGGCGGAGGAGGCGGCGG 2054  
 : : : ||||| : : : ||||| : : : ||||| ||  
 Db 493 -----ProGlyProAlaProHisGlyProProAlaProTyrAlaProGlnGlyTrpGly 510  
 QY 2055 GAAGTCTAGCCAGGTTTCCAGAA-----CCACGAGCCCGCC 2093  
 : : : ||||| : : : ||||| : : : ||||| ||  
 Db 511 AsnAlaTyrProHisTrpGlnGlnGlnAlaProProAspProAla 525

RESULT 14

PCT-US94-01782-10  
 ; Sequence 10, Application PC/TUS9401782  
 ; GENERAL INFORMATION:  
 ; APPLICANT: THE GOVERNMENT OF THE UNITED STATES  
 ; APPLICANT: AS REPRESENTED BY THE SECRETARY, DEPARTMENT OF  
 ; APPLICANT: HEALTH AND HUMAN SERVICES  
 ; TITLE OF INVENTION: NOVEL FUSE BINDING  
 ; TITLE OF INVENTION: PROTEIN AND CDNA THEREFOR  
 ; NUMBER OF SEQUENCES: 10  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: MORGAN & FINNEGAN  
 ; STREET: 345 PARK AVENUE  
 ; CITY: NEW YORK  
 ; STATE: NEW YORK  
 ; COUNTRY: USA



Db 367 pAsnMetGlyProProGlyGlyLeuGlnGluPheAsnPheIle----- 381  
QY 1593 GNAGCTGAGCGCATATCAGAGTGCCTCTCCAGAGCTGCCGGGTGATTGGCAAGG 1652  
Db 382 -----ValProThrGlyLysThrGlyLeuIleIleGlyLysG 394  
QY 1653 TGGCAAGACCGTGAACGAACCTTACCAGTGCAGAGTGCATCGTCCCTCGTGA 1712  
Db 394 yGlyGluThrIleLysSerGlnGlnSerGlyAlaArgIleGluLeuGlnArgAs 414  
QY 1713 CCAAGCGCAGATGAAATGAGAAAGTG---ATCGTCAGAATATTCGGGCAC----- 1761  
Db 414 nProProProAsnAlaAsProAsnMetLysLeuPheThrIleArgGlyThrProGlnG 434  
QY 1762 ----TTCITTTGTCAGCACTGCACAGCGCAAGATCAGGAAATGTACACAGGTGAA 1817  
Db 434 nIleAspTyrAlaArgGlnLeuIleGluLysIleGlyProValAsnProLeuG 454  
QY 1818 GCAGCAGGAGCAGAAATACCTCAGGAGTGCCTCACAGCGCAGCAAGTGAGGTCCCA 1877  
Db 454 yProProValProHisGlyProHisGlyVal-ProGlyProHisGlyProGlyProp 474  
QY 1878 CAGGCACCGCAGAAACACCGAGTGAATGTAGCCCTTCCAAACCTGCACAGATGACCA 1937  
Db 474 rGly-ProGly-----ThrPro-----MetGlyPro 482  
QY 1938 AACGCAGCGCAGATCGGAGCAACCAAGACCATCTGAGGAATGAGAAGTCTGCGG 1997  
Db 483 TyrAsnProAlaProTyrAsnProGlyPro----- 492  
QY 1998 AGCGCGCCAGGAGTCTGCGGAGG---CCCTGAGAACCCAGGGCCGAGGAGGCGGG 2054  
Db 493 -----ProGlyProAlaProHisGlyProProAlaProTyrAlaProGlnGlyTrpGly 510  
QY 2055 GAAGTGCAGCGAGTTTGCAGAA-----CCACGAGCCCGCC 2093  
Db 511 AsnAlaTyrProHisTrpGlnGlnAlaProProAspProAla 525

RESULT 15

US-08-021-608D-8  
Sequence 8, Application US/08021608D  
Patent No. 5580760  
GENERAL INFORMATION:  
APPLICANT: LEVENS, DAVID L., DUNCAN,  
APPLICANT: ROBERT C., AND AVIGAN, MARK I.  
TITLE OF INVENTION: NOVEL FUSE BINDING  
TITLE OF INVENTION: PROTEIN AND CDNA THEREFOR  
NUMBER OF SEQUENCES: 24  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MORGAN & FINNEGAN  
STREET: 345 PARK AVENUE  
CITY: NEW YORK  
STATE: NEW YORK  
COUNTRY: USA  
ZIP: 10154  
COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY DISK  
COMPUTER: IBM PC COMPATIBLE  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WORDPERFECT 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/021,608D  
FILING DATE: 22-FEB-1993  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: WILLIAM S. FEILER  
REGISTRATION NUMBER: 26,728  
REFERENCE/DOCKET NUMBER: 2026-4063  
TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 758-4800  
TELEFAX: (212) 751-6849  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 590  
TYPE: Amino Acid  
STRANDEDNESS: Single  
TOPOLOGY: Unknown  
MOLECULE TYPE: Peptide/Protein  
HYPOTHETICAL: No  
ORIGINAL SOURCE:  
ORGANISM: Human  
CELL LINE: HL60  
US-08-021-608D-8  
Alignment Scores:  
Pred. No.: 3,48e-12 Length: 590  
Score: 236.50 Matches: 121  
Percent Similarity: 35.93% Conservative: 68  
Best Local Similarity: 23.00% Mismatches: 182  
Query Match: 3,86% Indels: 155  
Gaps: 23  
DB:  
US-09-270-437D-6 (1-3412) x US-08-021-608D-8 (1-590)  
QY 667 GTCCCCACCCAGCTTTGTGGTCCCATCATCGAAAGAGGGCTTGACCAATAAAGAACATC 726  
Db 91 ValProAspGlyMetValGlyPheIleIleGlyArgGlyGlyGluGlnIleSerArgIle 110  
QY 727 ACTAAGCAGACCCAGTCCCGGTAGATATCCATAGAAAGAGAATCTGTGAGCTGCAGAG 786  
Db 111 GlnGlnGluSerGlyCysLysIleGlnIle---AlaProAspSerGlyGlyLeuProGlu 129  
QY 787 AGCCTGTCCACCATCCATGCCACCCAGAGGGGACTTCTGAAGCATGCCGCGCATGTTCTT 846  
Db 130 ArgSerCysMetLeuThrGlyThrProGluSerValGlnSerAlaLysArgLeuLeuAsp 149  
QY 847 GAATCATGCAGAAA-----GAGGCAGATGAGACCAACTA 882  
Db 150 GlnIleValGluLysGlyArgProAlaProGlyPheHisHisGlyAspGlyProGlyAsn 169  
QY 883 GCCGAAGAGATTCCTCTGAAATCTTGGCACAACAATGGCTTGGTTGGAAAGACTGATTGGA 942  
Db 170 AlaValGln-----GluIleMetIleProAlaSerLysAlaGlyLeuValIleGly 186  
QY 943 AAAGAAGCAGAAATTTGAGAAATTTGAACATGAACAGGAGCCAGAGATAACAATCTCA 1002  
Db 187 LysGlyGlyGluThrIleLysGlnLeuGlnGluArgAlaGlyValLysMetValMet--- 205  
QY 1003 TCTTTGCGAGGATTTGAGCATATATACACCGGAAAGAACCATCACTGTGAAGGGCACAGTT 1062  
Db 206 ---IleGlnAsp-----GlyProGlnAsnThr----- 213  
QY 1063 GAGGCTGTGCGAGTGTGAGATGAGATTAAGAAGAGCTGCTGGTGGAGCCCTTTGAAAT 1122  
Db 214 -----GlyAlaAsp-----LysProLeuArg-----IleThrGly 223  
QY 1123 GATATGCTGGCTTTAAACCAACAGCAATCTGATCCAGGGTGTGAACCTCAGCGCACTT 1182  
Db 224 AspProTyrLysValGlnGlnAlaLysGluMetVal----- 235  
QY 1183 GGCATCTTTTCAACAGGAGTGTCCGTGTATCTCCACCAGAGGGCCCGCGAGTCC 1242  
Db 236 -----LeuGluLeuIleArgAspGlnGlyGlyPheArgGluVal--- 248  
QY 1243 CCCGCTGCCCCCTACACCCCTTCACTACCCACTCGGATCTTCCAGCCTGTACCCC 1302  
Db 248 ----- 248  
QY 1303 CATCACCAGTTTGGCCCGTTCCCGCATCATCACTCTTATCCAGAGCAGAGATGTGAAT 1362  
Db 249 ArgAsnGluTyrGly-----SerArgIleGlyGlyAsnGluGlyIleAsp 263



```
QY 1363 CTCTTCATCCCAACCCAGGCTGTGGCCCATCATCGGGAAGAAGGGGGCACACATCAAA 1422
Db   :: ||||| ||||| ||||| ||||| ::||| ||||| ||||| |||||
264 ValProIleProArgPheAlaValGlyLeuValIleGlyArgAsnGlyGluMetIleLys 283
QY 1423 CAGTCGCGAGATTGCGGAGCCCTCTATCAAGATTGCCCTCGGGAAGGCCAGACGTC 1482
Db   ::||| ::||| ::||| ::||| ::||| ::||| ::||| ::|||
284 LysIleGlnAsnAspAlaGlyValArgIleGlnPheLysProAspAspGly---ThrThr 302
QY 1483 AGCGAAAGGATGTCATCATCACCGGGCCACCGGAA---GCCAGTTCAAGGCCAG--- 1536
Db   ||||| ::||| ||||| ||||| ||||| ||||| ||||| |||||
303 ProGluArgIleAlaGlnIleThrGlyProProAspArgCysGlnHisAlaAlaGluIle 322
QY 1537 -----GGG 1539
Db   |||
323 IleThrAspLeuLeuArgSerValGlnAlaGlyAsnProGlyGlyProGlyGlyGly 342
QY 1540 CGGATCTTTGGGAAA-----CTGAAAGAG 1563
Db   ||||| ::||| ::||| ::||| ::||| ::||| ::||| ::|||
343 ArgGlyArgGlyArgGlyGlnGlyAsnTrpAsnMetGlyProProGlyGlyLeuGlnGlu 362
QY 1564 GAAACTTCTTTAAACCCCAAGAGAGAGTGAACTGGAAGCGCATATCAGAGTCCCTCT 1623
Db   ||||| ||||| ||||| ||||| ||||| ||||| |||||
363 PheAsnPheIle-----ValProThr 369
QY 1624 TCCACAGCTGCGCGGTGATTCGCAAGGTGGCAAGCCGTGAACGAACTCCAGAACTTA 1683
Db   ||||| ::||| ||||| ||||| ||||| ||||| ||||| |||||
370 GlyLysThrGlyLeuIleIleGlyLysGlyGlyGluThrIleLysSerIleSerGlnGln 389
QY 1684 ACCAGTCGCAAGATCATGTCCTCCGTGACCAACAGCGCCAGATGAAATAGGAAAGTG--- 1740
Db   ::||| ::||| ::||| ::||| ::||| ::||| ::||| ::|||
390 SerGlyAlaArgIleGluLeuGlnArgAsnProProAsnAlaAspProAsnMetLys 409
QY 1741 ATCCTCAGATATTCGGGCAC-----TTCCTTTGTAGCCAGCTGCACAGCGC 1788
Db   ::||| ||||| ||||| ||||| ||||| ||||| |||||
410 LeuPheThrIleArgGlyThrProGlnGlnIleAspTyrAlaArgGlnLeuIleGluGlu 429
QY 1789 AAGATCAGGGAAATTGTACACAGGTGAAGCAGCAGGAGCAGAAATACCTCAGGGAGTC 1848
Db   ||||| ||||| ||||| ::||| ::||| ||||| |||||
430 LysIleGlyGlyProValAsnProLeuGlyProProValProHisGlyProHisGlyVal 449
QY 1849 GCCTCAGCGCAGCAAGTGAGGCTCCACAGCGCAGCAGCAAAACACGGATGAATGTAG 1908
Db   ||||| ||||| ||||| ||||| |||||
450 -ProGlyProHisGlyProProGlyProProGly-ProGly----- 462
QY 1909 CCCTTCCACACCTGACAGATGAGACCAACCGCAGCCAGCCAGATCGGGAGCAACCAA 1968
Db   ||||| ||||| ||||| ::||| ::||| ||||| |||||
463 -----ThrPro-----MetGlyProTyrAsnProAlaProTyrAsnProGlyPro- 477
QY 1969 AGACCATCTGAGGAATGAGAAGTCTCGGAGCGCGCCAGGACTCTGCCGAGG---CCCT 2025
Db   ||||| ||||| ||||| ||||| ||||| |||||
478 -----ProGlyProAlaProHisGlyProp 486
QY 2026 GAGAACCCCGGGCGGAGGGCGGGGAAGGTCTAGCCAGGTTTGCAGAA-----C 2079
Db   ||||| ||||| ::||| ||||| ||||| |||||
486 roAlaProTyrAlaProGlnGlyTrpGlyAsnAlaTyrProHisTrpGlnGlnGlnAlaP 506
QY 2080 CACCGAGCCCGCC 2093
Db   ||||| ||||| |||||
506 roProAspProAla 510
```

Search completed: July 16, 2004, 10:47:06  
Job time : 83 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: July 16, 2004, 10:33:46 ; Search time 318 Seconds  
(without alignments)  
6707.276 Million cell updates/sec

Title: US-09-270-437D-6  
Perfect score: 6121  
Sequence: 1 ggcagcgaggagcgagga.....aaccttgaaaattttattt 3412

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 1285345 seqs, 312560633 residues  
Total number of hits satisfying chosen parameters: 2570690  
Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ n2p.spool\_p/US09270437/runat 16072004 113051 13449/app query.fasta\_1.3591  
-Q=/cgn2\_1/USPTO.spool\_p/US09270437/runat 16072004 113051 13449  
-DB=Published Applications AA -QFMT=fastan -SUFFIX=rapb -MINMATCH=0.1  
-LOOPCL=0 -LOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blcosum62  
-TRANS-human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct-THR MAX=100  
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HAPSIZ=500 -MINLEN=0  
-MAXLEN=200000000 -USER=US09270437 @CGN 1 1 250 @runat 16072004 113051 13449  
-NCPU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100  
-LONGLOG -DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications\_AA.\*  
1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*  
17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	DB ID	Description
-----						

1	3065.5	50.1	620	9	US-09-764-864-1116	Sequence 1116, Ap
2	3003	49.1	587	15	US-10-313-986-501	Sequence 501, App
3	2806.5	45.9	556	14	US-10-097-340-147	Sequence 147, App
4	2806.5	45.9	556	16	US-10-648-593-182	Sequence 182, App
5	2553	41.7	555	15	US-10-262-445-40	Sequence 40, Appl
6	2025.5	33.1	577	15	US-10-313-986-500	Sequence 500, App
7	2019.5	33.0	577	9	US-09-873-637-2	Sequence 2, Appli
8	1968.5	32.2	589	15	US-10-313-986-486	Sequence 486, App
9	1957.5	32.0	579	9	US-09-735-705-348	Sequence 348, App
10	1957.5	32.0	579	9	US-09-850-716A-348	Sequence 348, App
11	1957.5	32.0	579	9	US-09-897-778-348	Sequence 348, App
12	1957.5	32.0	579	9	US-09-897-778-446	Sequence 446, App
13	1957.5	32.0	579	9	US-09-897-778-449	Sequence 449, App
14	1957.5	32.0	579	12	US-10-007-700-348	Sequence 348, App
15	1957.5	32.0	579	12	US-10-007-700-446	Sequence 446, App
16	1957.5	32.0	579	12	US-10-007-700-449	Sequence 449, App
17	1957.5	32.0	579	14	US-10-117-982-348	Sequence 348, App
18	1957.5	32.0	579	14	US-10-117-982-446	Sequence 446, App
19	1957.5	32.0	579	14	US-10-117-982-449	Sequence 449, App
20	1957.5	32.0	579	14	US-10-117-982-480	Sequence 480, App
21	1957.5	32.0	579	15	US-10-313-986-348	Sequence 348, App
22	1957.5	32.0	579	15	US-10-313-986-446	Sequence 446, App
23	1957.5	32.0	579	15	US-10-313-986-449	Sequence 449, App
24	1957.5	32.0	579	15	US-10-313-986-480	Sequence 480, App
25	1952.5	31.9	586	9	US-09-850-716A-427	Sequence 427, App
26	1952.5	31.9	586	9	US-09-897-778-427	Sequence 427, App
27	1952.5	31.9	586	14	US-10-007-700-427	Sequence 427, App
28	1952.5	31.9	586	14	US-10-117-982-427	Sequence 427, App
29	1952.5	31.9	586	15	US-10-313-986-427	Sequence 427, App
30	1950.5	31.9	579	9	US-09-735-705-176	Sequence 176, App
31	1950.5	31.9	579	9	US-09-850-716A-176	Sequence 176, App
32	1950.5	31.9	579	9	US-09-897-778-176	Sequence 176, App
33	1950.5	31.9	579	10	US-09-466-396A-176	Sequence 176, App
34	1950.5	31.9	579	12	US-10-007-700-176	Sequence 176, App
35	1950.5	31.9	579	14	US-10-117-982-176	Sequence 176, App
36	1950.5	31.9	579	15	US-10-313-986-176	Sequence 176, App
37	1947.5	31.8	579	14	US-10-117-982-484	Sequence 484, App
38	1947.5	31.8	579	15	US-10-313-986-484	Sequence 484, App
39	1408.5	23.0	432	16	US-10-408-765A-2088	Sequence 2088, Ap
40	910	14.9	192	9	US-09-764-864-1117	Sequence 1117, Ap
41	898	14.7	261	9	US-09-764-864-1114	Sequence 1114, Ap
42	868	14.2	250	9	US-09-764-864-1532	Sequence 1532, Ap
43	803	13.1	171	9	US-09-764-864-1119	Sequence 1119, Ap
44	750	12.3	171	9	US-09-764-864-1536	Sequence 1536, Ap
45	357	5.8	81	14	US-10-117-982-476	Sequence 476, App

ALIGNMENTS

RESULT 1  
US-09-764-864-1116  
; Sequence 1116, Application US/09764864  
; Patent No. US20020132753A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PT223  
; CURRENT APPLICATION NUMBER: US/09/764,864  
; CURRENT FILING DATE: 2001-01-17  
; Prior application data removed - consult PALM or file wrapper  
; NUMBER OF SEQ ID NOS: 1792  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1116  
; LENGTH: 620  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SITE  
; LOCATION: (533)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
US-09-764-864-1116

Alignment Scores:

Pred. No.: 1,01e-232 Length: 620  
 Score: 3065.50 Matches: 603  
 Percent Similarity: 97.73% Mismatches: 1  
 Best Local Similarity: 97.57% Indels: 11  
 Query Match: 50.08% Gaps: 3  
 DB: 9

US-09-270-437D-6 (1-3412) x US-09-764-864-1116 (1-620)

QY	22	CCCGGGTACCGGGCGGGGAGCGGGCTCTCGGGGAAGAGACGG-----ATG	72
Db	3	ArgArgTyrAlaCysArgTyrArgSerGlyIleProGlySerThrHisAlaSerGlyMet	22
QY	73	ATGAACAAGCTTTACATCGGAACCTTGAGCCCGCGCTCACCGCGAGACCTCGGGGAG	132
Db	23	MetAsnLysLeuTyrIleGlyAsnLeuSerProAlaValThrAlaAspLeuArgGln	42
QY	133	CTCTTTGGGGACAGGAAGCTGCCCTGGCGGACAGGTCCTGCTGAAGTCCGGCTACGCC	192
Db	43	LeuPheGlyAspArgLysLeuProLeuAlaGlyGlnValLeuLeuLysSerGlyTyrAla	62
QY	193	TTCTGTTGACTACCCGACAGACTGGGCCATCCGGCCATCGAGACCTCTCGGGTAAA	252
Db	63	PheValAspTyrProAspGlnAsnTrpAlaIleArgAlaIleGluThrLeuSerGlyLys	82
QY	253	GTGGAATTGCTGGGAAATCATGGAAGTTGATTACTCAGTCTCTAAAAAGCTAAGGAGC	312
Db	83	ValGluLeuHisGlyLysIleMetGluValAspTyrSerValSerLysLysLeuArgSer	102
QY	313	AGGAAATTCAGATTGGAACATCCCTCCTACCTGCAGTGGGAGGTGTGGATGACTT	372
Db	103	ArgLysIleGlnIleArgAsnIleProProHisLeuGlnTrpGluValLeuAspGlyLeu	122
QY	373	TTGGCTCAATATGGACAGCTGAGATGTGGAACTGGAACAGTCAACACACAGAAACCGCC	432
Db	123	LeuAlaGlnTyrGlyThrValGluAsnValGluGlnValAsnThrAspThrGluThrAla	142
QY	433	GTGTCAACGTCACATATGCAACAAGAGAAGCAAAATAGCCATGGAGAAGCTAAGC	492
Db	143	ValValAsnValThrTyrAlaThrArgGluAlaLysIleAlaMetGluLysLeuSer	162
QY	493	GGGATCATGTTGAGAACTACTCTTCAAGATTCTCATATCCCGATCAAGAGTGTGAC	552
Db	163	GlyHisGlnPheGluAsnTyrSerPheLysIleSerTyrIleProAspGluValSer	182
QY	553	TCCCTTCGCCCTCAGCGACCCAGCTGGGGACCACTCTTCCGGGAGCAAGCCAC	612
Db	183	SerProSerProGlnArgAlaGlnArgGlyAspHisSerSerArgGluGlnGlyHis	202
QY	613	GCCCTGGGGGACATTCTCAGGCCACAGAGATTGATTTCCCGCTCGCGATCCTGTCGCC	672
Db	203	AlaProGlyGlyThrSerGlnAlaArgGlnIleAspPheProLeuArgIleLeuValPro	222
QY	673	ACCGATTGTTGGTGCATCATCGAAAGAGGCTTGACATAAAGAACATCACTAAG	732
Db	223	ThrGlnPheValGlyAlaIleIleGlyLysGluGlyLeuThrIleLysAsnIleThrLys	242
QY	733	CAGACCCAGTCCCGGTAGATATCCATAGAAAGAGAACTCTGGAGCTGCAGAGAGCT	792
Db	243	GlnThrGlnSerArgValAspIleHisArgLysGluAsnSerGlyAlaAlaGlnLysPro	262
QY	793	GTCCACCATCCATGCCACCCGAGGGGACTTCTGAAGCATGCCGATGATCTTGAATC	852
Db	263	ValThrIleHisAlaThrProGluGlyThrSerGluAlaCysArgMetIleLeuGluIle	282
QY	853	ATGAGAAAGGCGAGTACAGCAACTAGCCGAGAGATCTCTGAAAATCTTGCA	912
Db	283	MetGlnLysGluAlaAspGluThrLysLeuAlaGluGluIleProLeuLysIleLeuAla	302
QY	913	CACATGCTTGGTGGAGACTGATTGGAAGAGGACAGAAATTTGAAGAAATTTGAA	972
Db	303	HisAsnGlyLeuValGlyArgLeuIleGlyLysGluGlyArgAsnLeuLysIleGlu	322

QY	973	CATGAACAGGACCAAGATAACAATCTCATCTTTGAGGATTTGAGCATATACACCGG	1032
Db	323	HisGluThrGlyThrLysIleThrIleSerSerLeuGlnAspLeuSerIleTyrAsnPro	342
QY	1033	CAAGAACAATCCTGTGAAAGGGCACAGTTGAGGCTGTGCCAGTGTGAGATAGAGATT	1092
Db	343	GluArgThrIleThrValLysGlyThrValGluAlaCysAlaSerAlaGluIleGluIle	362
QY	1093	ATGAAGAAGCTGCTGAGGCTTTGAAAATGATATGCTGGCTTTAAACAACAGCCAAAT	1152
Db	363	MetLysLysLeuArgGluAlaPheGluAsnAspMetLeuAlaValAsnGlnGlnAlaAsn	382
QY	1153	CTGATCCCGAGGTGAACCTCAGCGACTGTGCATCTTTCAACAGGACTGTCCTGCTA	1212
Db	383	LeuIleProGlyLeuAsnLeuSerAlaLeuGlyIlePheSerThrGlyLeuSerValLeu	402
QY	1213	TCTCCACGACGAGGCGCGGAGCTCCCGCTGCGCCCTACACCCCTTCTACTACC	1272
Db	403	SerProProAlaGlyProArgGlyAlaProProAlaAlaProTyrHisProPheThr	422
QY	1273	CATCCCGATATCTTCTCAGCGCTGTACCCCATCACAGTTTGGCCCGTTCGCCGATCAT	1332
Db	423	HisSerGlyTyrPheSerSerLeuTyrProHisHisGlnPheGlyProPheProHisHis	442
QY	1333	CACCTCTATCCAGACGACGAGATTGTGAATCTCTTATCCCAACCCAGGCTGTGGCGCC	1392
Db	443	HisSerTyrProGluGlnGluIleValAsnLeuPheIleProThrGlnAlaValGlyAla	462
QY	1393	ATCATCGGGAAGAGGGGGCACACATCAACAGCTGGGAGATTTCGCGGAGCCCTCTATC	1452
Db	463	IleIleGlyLysLysGlyAlaHisIleLysGlnLeuAlaArgPheAlaGlyAlaSerIle	482
QY	1453	AAGATTGCCCTTCGCGAAGGCCAGACGTCAGCGAAAGATGTGTATCATCACCGGGCCA	1512
Db	483	LysIleAlaProAlaGluGlyProAspValSerGluArgMetValIleIleThrGlyPro	502
QY	1513	CCGGAAGCCCATGTTCAAGGCCGAGGAGCGGATCTTTGGAACTCAAAAGAGAAATTC	1572
Db	503	ProGluAlaGlnPheLysAlaGlnGlyArgIlePheGlyLysLeuLysGluGluAsnPhe	522
QY	1573	TTTAAACCCCAAGAAAGTGAAGCTGGAAGCGCATATCAGAGTGCCTCTTCCACAGCT	1632
Db	523	PheAsnProLysGluGluValLysLeuGlu**HisIleArgValProSerSerThrAla	542
QY	1633	GGCCGGTGTATGGCAAGCTGGACAGCCGTGACAGCACTGACAGCACTTAACAGTGA	1692
Db	543	GlyArgValIleGlyLysGlyLysThrValAsnGluLeuGlnAsnLeuThrSerAla	562
QY	1693	GAAGTCATCGTCCCTCGTGACCAACGCCAGATGAAATGAGGAAGTGCATCGTCAGAATT	1752
Db	563	GluValIleValProArgAspGlnThrProAspGluAsnGluGluValIleValArgIle	582
QY	1753	ATCGGCACTTCTTTGCTAGCCAGCTGCACAGCGCAAGATCAGGGAATTTGTACAAAG	1812
Db	583	IleGlyHisPhePheAlaSerGlnThrAlaGlnArgLysIleArgGluIleValGlnGln	602
QY	1813	GTGAAGCAGCAGGACGAGAATACCTCAGGAGTGCCTCACAGCCGAGCAAG	1866
Db	603	VallysGlnGlnGluGlnLysTyrProGlnGlyValAlaSerGlnArgSerLys	620

RESULT 2  
 US-10-313-986-501  
 ; Sequence 501, Application US/10313986  
 ; Publication No. US20030236209A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Foy, Teresa M.  
 ; APPLICANT: McNabb, Andria  
 ; APPLICANT: Watanabe, Yoshihiro  
 ; APPLICANT: Reed, Steven G.  
 ; APPLICANT: Wang, Tongtong  
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
 ; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER  
 ; FILE REFERENCE: 210121.455C19

; CURRENT APPLICATION NUMBER: US/10/313,986  
; CURRENT FILING DATE: 2002-12-04  
; NUMBER OF SEQ ID NOS: 560  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 501  
; LENGTH: 587  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-313-986-501

## Alignment Scores:

Pred. No.: 8,496-228 Length: 587  
Score: 2003.00 Matches: 587  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 49.06% Indels: 0  
DB: 15 Gaps: 0

US-09-270-437D-6 (1-3412) x US-10-313-986-501 (1-587)

QY	73	ATGAACAAGCTTTACATCGGGAACCTGAGCCCGCGCTACCGCGCAGCAGCTCCGCGCAG	132
Db	1	MetAsnLysLeuTyrlleGlyAsnLeuSerProAlaValThrAlaAspAspLeuArgGln	20
QY	133	CTCTTTGGGACACGAGAGCTGCCCTCGCGGACAGGTCCTGCTGAAGTCGGCTACGCC	192
Db	21	LeuPheGlyAspArgLysLeuProLeuAlaGlyGlnValLeuLeuLysSerGlyTyAla	40
QY	193	TTGCTGGACTACCCCGACACAGAACTGGGCCATCGCGCCATCGAGACCCCTCTCGGTA	252
Db	41	PheValAspTyrProAspGlnAsnTrpAlaIleArgAlaIleGluThrLeuSerGlyLys	60
QY	253	GTGGAATTGATGGGAAAATCATGGAAGTTGATTACTCAGTCTCTAAAAAGCTAAGGAGC	312
Db	61	ValGluLeuHisGlyLysIleMetGluValAspTyrSerValSerLysLysLeuArgSer	80
QY	313	AGGAATTCAGATTGCAATACCTCTCCTCAGCTGAGTGGGAGGTGTTGGATGACCTT	372
Db	81	ArgLysIleGlnIleArgAsnIleProHisLeuGlnTrpGluValLeuAspGlyLeu	100
QY	373	TTGGCTCAATATGGGACAGTGGGAATGTGGAACAGTCAACACAGACACAGAAACCGCC	432
Db	101	LeuAlaGlnTyrGlyThrValGluAsnValGluGlnValAsnThrAspThrGluThrAla	120
QY	433	GTGTGCAAGTTCATATGCAACAGAGAGAAAGCAAAAATAGCCATGGAGAACTAAGC	492
Db	121	ValValAsnValThrTyrAlaThrArgGluGluAlaLysIleAlaMetGluLysLeuSer	140
QY	493	GGGCATCAGTTTGAGAACTACTCTTCAAGATTTCTACATCCCGGATGAGAGGTGAGC	552
Db	141	GlyHisGlnPheGluAsnTyrSerPheLysIleSerTyrIleProAspGluGluValSer	160
QY	553	TCCCTTTCGCCCTCAGCGAGCCGAGCGTGGGACCACTCTTCCCGGAGCAAGGCCAC	612
Db	161	SerProSerProGlnArgAlaGlnArgGlyAspHisSerSerArgGluGlnGlyHis	180
QY	613	GCCCTCGGGCACTTCTCAGCCAGACAGATTTCCCTGCGTCCGATCCCTGCTCCCTCCC	672
Db	181	AlaProGlyGlyThrSerGlnAlaArgGlnIleAspPheProLeuArgIleLeuValPro	200
QY	673	ACCCAGTTGTTGCTGCCATCATCGAAAGAGGGCTTGACCATAAAGACATCATCAAG	732
Db	201	ThrGlnPheValGlyAlaIleIleGlyLysGluGlyLeuThrIleLysAsnIleThrLys	220
QY	733	CAGACCCAGTCCCGGTAGATATCCATAGAAAAGAGAACTCTGGAGCTGCAGAGAGCCT	792
Db	221	GlnThrGlnSerArgValAspIleHisArgLysGluAsnSerGlyAlaAlaGluLysPro	240
QY	793	GTCAACATCCATGCCACCCAGAGGGGACTTCTGAAGCATCCCGCATGATTCTTGAATC	852
Db	241	ValThrIleHisAlaThrProGluGlyThrSerGluAlaCysArgMetIleLeuGluIle	260
QY	853	ATGCAGAAAGAGGCGATGATGAGACCAACTAGCCGAGAGGATCTCTGAAATCTTGCCA	912

Db	261	MetGlnLysGluAlaAspGluThrLysLeuAlaGluGluLeuProLeuLysIleLeuAla	280
QY	913	CACAATGGCTTGGTGAAGACTGATTGGAAGAAAGAGGAGAAATTTGAAGAAATGAA	972
Db	281	HisAsnGlyLeuValGlyArgLeuIleGlyLysGluGlyArgAsnLeuLysLysIleGlu	300
QY	973	CATGAACAGGGACCAAGATAACAATCTCATCTTTGAGGATTGTGAGCATATCAACCCG	1032
Db	301	HisGluThrGlyThrLysIleThrIleSerSerLeuGlnAspLeuSerIleTyrAsnPro	320
QY	1033	GAAGAACCATCACTGTGAAGGCGACAGTTGAGCCCTGTGCCAGTGTGTGAGATAGATT	1092
Db	321	GluArgThrIleThrValLysGlyThrValGluAlaCysAlaSerAlaGluIleGluIle	340
QY	1093	ATGAAGAAGTCGCTGAGGCTTTGAAATGATATGCTGCTGTAAACCAACAAGCCAAT	1152
Db	341	MetLysLysLeuArgGluAlaPheGluAsnAspMetLeuAlaValAsnGlnAlaAsn	360
QY	1153	CTGATCCAGGGTTGAACCTCAGGCACTTGGGCACTTTTCAACAGGACTGTCCGTGCTA	1212
Db	361	LeuIleProGlyLeuAsnLeuSerAlaLeuGlyIlePheSerThrGlyLeuSerValLeu	380
QY	1213	TCTCCACCCAGCGCCCGCGAGCTCCCGCGCTGCCCGCTACCAACCCCTTCACTACC	1272
Db	381	SerProProAlaGlyProArgGlyAlaProProAlaAlaProTyrHisProPheThrThr	400
QY	1273	CACCTCGGATPACTCTCCAGCTGTACCCCATCACAGTTTGGCCCGTTCGCGCATCAT	1332
Db	401	HisSerGlyTyrPheSerSerLeuTyrProHisGlnPheGlyProPheProHisHis	420
QY	1333	CACCTTATCCAGCAGGAGATTGGAATCTCTTATCCCAACCCAGGCTGTGGGCGCC	1392
Db	421	HisSerTyrProGluGlnGluIleValAsnLeuPheIleProThrGlnAlaValGlyAla	440
QY	1393	ATCATCGGAAGAGGGGCGACATCAACAGCTGCGAGATTGCGCGAGGCTCTATC	1452
Db	441	IleIleGlyLysGlyAlaHisIleLysGlnLeuAlaArgPheAlaGlyAlaSerIle	460
QY	1453	AAGATTGCCCTCGGGAAGCCCGAGCTCAGCGAAAGGATGGTCATCATCACCGGCCA	1512
Db	461	LysIleAlaProAlaGluGlyProAspValSerGluArgMetValIleIleThrGlyPro	480
QY	1513	CCGGAAGCCAGTTCAAGGCCCGGACGGATCTTGGGAAACTGAAAGAGGAAACTTC	1572
Db	481	ProGluAlaGlnPheLysAlaGlnGlyArgIlePheGlyLysLeuLysGluGluAsnPhe	500
QY	1573	TTTAAACCCCAAGAAAGTGAAGCTGGAAGCCCATATCAGAGTGCCTCTTCCACAGCT	1632
Db	501	PheAsnProLysGluGluValLysLeuGluAlaHisIleArgValProSerSerThrAla	520
QY	1633	GGCGGGTGATTGGCAAGGTGGCAGACCGGTGAACCGAAGTGAACCTTAAACAGTGCA	1692
Db	521	GlyArgValIleGlyLysGlyLysThrValAsnGluLeuGlnAsnLeuThrSerAla	540
QY	1693	GAAGTCACTGCTGTCACCAACCGCAGATGAAATCAGGAAGTGTGTCAGAAAT	1752
Db	541	GluValIleValProArgAspGlnThrProAspGluAsnGluGluValIleValArgIle	560
QY	1753	ATCGGCGACTTCTTTGCTAGCAGACTGCACGCAAGATCAGGGAATTTGTACACAG	1812
Db	561	IleGlyHisPheAlaSerGlnThrAlaGlnArgLysIleArgGluIleValGlnGln	580
QY	1813	GTGAAGCAGCAGGACAGAAA 1833	
Db	581	VallysGlnGlnGluGlnLys 587	

## RESULT 3

US-10-097-340-147  
; Sequence 147, Application US/10097340  
; Publication No. US20030087250A1  
; GENERAL INFORMATION:  
; APPLICANT: John MONAHAN

```

; APPLICANT: Manjula GANNAVARAPU
; APPLICANT: Sebastian HOERSCH
; APPLICANT: Shubhangi KAMATKAR
; APPLICANT: Steve G. KOVATS
; APPLICANT: Rachel E. MEYERS
; APPLICANT: Michael MORRISSEY
; APPLICANT: Peter OLANDT
; APPLICANT: Ami SEN
; APPLICANT: Peter VEIBY
; APPLICANT: Gordon B. MILLS
; APPLICANT: Robert C. EAST, Jr.
; APPLICANT: Karen LU
; APPLICANT: Rosemarie SCHMANDT
; APPLICANT: Xumei ZHAO
; APPLICANT: Karen GLATT
; TITLE OF INVENTION: Nucleic Acid Molecules and Proteins For The Identification,
; TITLE OF INVENTION: Assessment, Prevention, and Therapy of Ovarian Cancer
; FILE REFERENCE: MRI-030
; CURRENT APPLICATION NUMBER: US/10/097,340
; CURRENT FILING DATE: 2002-03-14
; PRIOR APPLICATION NUMBER: 60/276,025
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/325,149
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/276,026
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/324,967
; PRIOR FILING DATE: 2001/09/26
; PRIOR APPLICATION NUMBER: 60/311,732
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/325,102
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/323,580
; PRIOR FILING DATE: 2001-09-19
; NUMBER OF SEQ ID NOS: 363
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 147
; LENGTH: 556
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-097-340-147

```

## Alignment Scores:

```

Pred. No.: 2,66e-212 Length: 556
Score: 2806.50 Matches: 556
Percent Similarity: 92.82% Conservative: 0
Best Local Similarity: 92.82% Mismatches: 0
Query Match: 45.85% Indels: 43
DB: 14 Gaps: 1

```

US-09-270-437D-6 (1-3412) x US-10-097-340-147 (1-556)

```

QY 70 ATGATGAACAGCTTTACATCGGGAACCTGAGCCCGCGGTCCACGCGGACGACTCCGG 129
Db 1 MetMetAsnLysLeuTyrlleGlyAsnLeuSerProAlaValThrAlaAspLeuArg 20
QY 130 CAGCTCTTTGGGGACAGGAAGTGCCTCGCGGGACAGGTCCTCTGAAGTCCGGCTAC 189
Db 21 GlnLeuPheGlyAspArgLysLeuProLeuAlaGlyGlnValLeuLeuLysSerGlyTyr 40
QY 190 GCCTTCGTGGACTACCCGACAGAACTGGGCGCAATCCGCGCAATCGAGACCCCTCTCGG 249
Db 41 AlaPheValAspTyrProAspGlnAsnTrpAlaIleArgAlaIleGluThrLeuSerGly 60
QY 250 AAAGTGAATTCATGGGAAATCATGGAAGTTGATTACTCAGTCTCTAAAAGCTAAGG 309
Db 61 LysValGluLeuHisGlyLysIleMetGluValAspTyrSerValSerLysLysLeuArg 80
QY 310 AGCAGGAAATTCAGATTCGAAACATCCCTCTCACTGCGAGTGGAGGTGTGGATGGA 369
Db 81 SerArgLysIleGlnIleArgAsnIleProProHisLeuGlnTrpGluValLeuAspGly 100
QY 370 CTTTGGCTCATATGGGACAGTGGAGATGTGGNAACAGTCAACACAGACACAGAACCC 429

```

```

Db 101 LeuLeuAlaGlnTyrGlyThrValGluAsnValGluGlnValAsnThrAspThrGluThr 120
QY 430 GCCGTTGTCAACGTCACATATGCAACAAGAGAGAGCAAAATATGACCATGAGAGCTTA 489
Db 121 AlaValValAsnValThrTyrAlaThrArgGluAlaLysIleAlaMetGluLysLeu 140
QY 490 AGCGGGCATCAGTTTGAGAACTACTCTTCAAGATTTCTACATCCCGATCAAGAGGTG 549
Db 141 SerGlyHisGlnPheGluAsnTyrSerPheLysIleSerTyrIleProAspGluVal 160
QY 550 AGCTCCCTTCGCCCCCTCAGCGAGCCCGAGGTGGGACCACTCTTCCCGGAGCAAGGC 609
Db 161 SerSerProSerProProGlnArgAlaGlnArgGlyAspHisSerSerArgGluGlnGly 180
QY 610 CACGCCCTTCGGGGCATTCTCAGCCACAGATTGATTTCCTCCCTCGGATCTGGTTC 669
Db 181 HisAlaProGlyGlyThrSerGluAlaArgGlnIleAspPheProLeuArgIleLeuVal 200
QY 670 CCCACCCAGTTTGTGGTCCCATCATCGAAAGAGGAGGCTTGACCATAAAGAACATCACT 729
Db 201 ProThrGlnPheValGlyAlaIleGlyLysGluGlyLeuThrIleLysAsnIleThr 220
QY 730 AAGCAGACCCAGTCCCGGTAGATATCCATAGAAAGAGAACTCTGGAGCTGCAGAGAAG 789
Db 221 LysGlnThrGlnSerArgValAspIleHisArgLysGluAsnSerGlyAlaAlaGluLys 240
QY 790 CTTGTACCATCCATCCACCCAGAGGGGACTTCTGAAGCATGCCGATGATCTTGA 849
Db 241 ProValThrIleHisAlaThrProGluGlyThrSerGluAlaCysArgMetIleLeuGlu 260
QY 850 ATCATGCAGAAAGAGGAGATGAGACCAACTAGCCGAGAGAGATCTCTGAAATCTTG 909
Db 261 IleMetGlnLysGluAlaAspGluThrLysLeuAlaGluGluIleProLeuLysIleLeu 280
QY 910 GCACAAATGCGCTTGGTGGAAAGTGTGGAAGAGAGGAGGAGAGGAGGAGGAGGAGG 969
Db 281 AlaHisAsnGlyLeuValGlyArgLeuIleGlyLysGluGlyArgAsnLeuLysIle 300
QY 970 GAACATGAACACAGGACCAAGATTAACATCTCATCTTTGTCAGGATTTGAGATATACAA 1029
Db 301 GluHisGluThrGlyThrLysIleThrIleSerSerLeuGlnAspLeuSerIleTyrAsn 320
QY 1030 CCGGAAAGAACATCCTCAAGGCGACAGTTCAGGCGCTGTCCAGAGCTGAGATAGAG 1089
Db 321 ProGluArgThrIleThrValLysGlyThrValGluAlaCysAlaSerAlaGluIleGlu 340
QY 1090 ATTATGAAGAGCTCGGTGAGGCGCTTTGAAAATGATATGCTGGCTGTTAACCAAGCC 1149
Db 341 IleMetLysLysLeuArgGluAlaPheGluAsnAspMetLeuAlaValAsn----- 357
QY 1150 AATCTGATCCAGGGTTGAACCTCAGCGCATCTTGGCATCTTTTCAACAGGACTGTCCGTG 1209
Db 357 ----- 357
QY 1210 CTATCTCCACGAGGCGCGCGGAGCTCCCCCGGTGCCCCCTACCAACCCCTTCACT 1269
Db 357 ----- 357
QY 1270 ACCCACTCCGATACTTCTCAGCCTGTACCCCATCACCAAGTTTGGCCCGTTCGCCGAT 1329
Db 358 ThrHisSerGlyTyrPheSerSerLeuTyrProHisGlnPheGlyProPheProHis 377
QY 1330 CATCACTCTTATCCAGAGCAGAGATTGTGAATCTTCTATCCCAACCCAGCGCTGTGGC 1389
Db 378 HisHisSerTyrProGluGlnGluIleValAsnLeuPheIleProThrGlnAlaValGly 397
QY 1390 GCCATCATCGGAAAGAGGGGCGCACATCAAAAGCTGGCGAGATTCCCGGAGCTCT 1449
Db 398 AlaIleIleGlyLysLysGlyAlaHisIleLysGlnLeuAlaArgPheAlaGlyAlaSer 417
QY 1450 ATCAAGATTGCCCCCTGGGAGAGGCCAGACCTCAGCGAAAGGATGGTTCATCATCCGGG 1509

```

Db 418 IleLysIleAlaProAlaGluGlyProAspValSerGluArgMetValIleIleThrGly 437  
QY 1510 CCACCGAGCCAGTTCAAGCCGAGGACGGATCTTTGGGAAACTGAAAGAGAAAAC 1569  
Db 438 ProProGluAlaGlnPheLysAlaGlnGlyArgIlePheGlyLysLeuLysGluGluAsn 457  
QY 1570 TTTCTTAACCCCAAGAGAGAGTGAAGCTGGAAGCGCATATCAGAGTCCCTCTTCCACA 1629  
Db 458 PhePheAsnProLysGluGluValLysLeuGluAlaHisIleArgValProSerSerThr 477  
QY 1630 GCTGCCGGGTGATTGGCAAGGTGGCAAGACCGTGAACGAACTGCAGAACTTAACCAAGT 1689  
Db 478 AlaGlyArgValIleGlyLysGlyGlyLysThrValAsnGluLeuGlnAsnLeuThrSer 497  
QY 1690 GCAGAGTCATCGTCCCTCGTACCAACAGCCAGATGAAATGAGGAAGTATGATCGTCA 1749  
Db 498 AlaGluValIleValProArgAspGlnThrProAspGluAsnGluGluValIleValArg 517  
QY 1750 ATTATCGGGCACTTCTTTGCTAGCCAGACTGCACAGCGCAAGATCAGGGAATTTGTACAA 1809  
Db 518 IleIleGlyHisPhePheAlaSerGlnThrAlaGlnArgLysIleArgGluIleValGln 537  
QY 1810 CAGTGAAGCAGCAGGACGAGAAATACCTCAGGAGTCGCTCACAGCGCAGCAAG 1866  
Db 538 GlnValLysGlnGlnGluGlnLysTyrProGlnGlyValAlaSerGlnArgSerLys 556

RESULT 4  
US-10-648-593-182  
; Sequence 182, Application US/10648593  
; Publication No. US20040106132A1  
; GENERAL INFORMATION:  
; APPLICANT: Bristol-Myers Squibb Company  
; TITLE OF INVENTION: IDENTIFICATION OF GENES FOR PREDICTING ACTIVITY OF COMPOUNDS THAT  
; TITLE OF INVENTION: INTERACT WITH AND/OR MODULATE PROTEIN TYROSINE KINASES AND/OR  
; TITLE OF INVENTION: PROTEIN TYROSINE KINASE PATHWAYS IN BREAST CELLS  
; FILE REFERENCE: D0273 NP  
; CURRENT APPLICATION NUMBER: US/10/648,593  
; PRIOR FILING DATE: 2003-08-26  
; PRIOR APPLICATION NUMBER: 60/406,385  
; PRIOR FILING DATE: 2002-08-27  
; NUMBER OF SEQ ID NOS: 557  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 182  
; LENGTH: 556  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-648-593-182

Alignment Scores:  
Pred. No.: 2,66e-212 Length: 556  
Score: 2806.50 Matches: 556  
Percent Similarity: 92.82% Conservative: 0  
Best Local Similarity: 92.82% Mismatches: 0  
Query Match: 45.85% Indels: 43  
DB: 16 Gaps: 1

US-09-270-437D-6 (1-3412) x US-10-648-593-182 (1-556)

QY 70 ATGATGAACAGCTTTTACATCGGAACCTGAGCCCGCGCTACCGCGCAGCAGCTCCGG 129  
Db 1 MetMetAsnLysLeuTyrIleGlyAsnLeuSerProAlaValThrAlaAspLeuArg 20  
QY 130 CAGCTCTTTGGGACAGGAAGCTCCCTGGGGGACAGGTCCTGCTGAAGTCGGGTAC 189  
Db 21 GlnLeuPheGlyAspArgLysLeuProLeuAlaGlyGlnValLeuLeuLysSerGlyTyr 40  
QY 190 GCCTTCGTGGACTACCCGACAGAACTGGGCCATCGCGCCATCGAGACCTCTCGGGT 249  
Db 41 AlaPheValAspTyrProAspGlnAsnTrpAlaIleArgAlaIleGluThrLeuSerGly 60  
QY 250 AAATGGAATTCATGGGAAATCATGGAAGTGTATTACTCAGTCTCTAAAGCTAAGG 309  
Db 61 LysValGlnLeuHisGlyLysIleMetGluValAspTyrSerValSerLysLysLeuArg 80

QY 310 AGCAGGAAAATTCAGATTGAAAACATCCCTCTCACTGSCAGTGGAGGTGTGGATGGA 369  
Db 81 SerArgLysIleGlnIleArgAsnIleProHisLeuGlnTrpGluValLeuAspGly 100  
QY 370 CTTTGGCTCAATATGGGACAGTGGAGNAATGTGAAACAAGTCAACACACACACAGAAACC 429  
Db 101 LeuLeuAlaGlnTyrGlyThrValGluAsnValGluGlnValAsnThrAspThrGluThr 120  
QY 430 GCCGTGTGCAACGTCACATATGCAACAGAGAAAGCAAAATAGCCATGGAGAAGCTA 489  
Db 121 AlaValValAsnValThrTyrAlaThrArgGluAlaLysIleAlaMetGluLysLeu 140  
QY 490 AGCGGGCATCAGTTTGAGAACTACTCTCTCAAGATTCTCTACATCCCGGATGAAGAGTG 549  
Db 141 SerGlyHisGlnPheGluAsnTyrSerPheLysIleSerTyrIleProAspGluGluVal 160  
QY 550 AGCTCCCTTCGCCCCCTCAGCGAGCCAGCGTGGGACCACTCTTCCGGGAGCAAGGC 609  
Db 161 SerSerProSerProGlnArgAlaGlnArgGlyAspHisSerSerArgGluGlnGly 180  
QY 610 CACGCCCTTGGGGGCATCTTCTCAGGCCAGACAGATTGATTTCCCGCTCGCGATCTCGTC 669  
Db 181 HisAlaProGlyGlyThrSerGlnAlaArgGlnIleAspPheProLeuArgIleLeuVal 200  
QY 670 CCCACCCAGTTTGTGTGCCATCATCGGAAAGAGGGCTTGACCATAAAGAACATCACT 729  
Db 201 ProThrGlnPheValGlyAlaIleGlyLysGluGlyLeuThrIleLysAsnIleThr 220  
QY 730 AAGCAGACCCAGTCCCGGTAGATATCCATAGAAAGAGACTCTGGAGCTGCAGAGAG 789  
Db 221 LysGlnThrGlnSerArgValAspIleHisArgLysGluAsnSerGlyAlaAlaGluLys 240  
QY 790 CCTCTCACCATCCATGCCACCCAGAGGGAGCTTCTGAAGCATCCGCGATGATTTCTGAA 849  
Db 241 ProValThrIleHisAlaThrProGluGlyThrSerGluAlaCysArgMetIleLeuGlu 260  
QY 850 ATCATGACAGAAAGGAGGATGAGACCAACTAGCCGAAAGAGATTCTCTGAAATCTTG 909  
Db 261 IleMetGlnLysGluAlaAspGluThrLysLeuAlaGluGluIleProLeuLysIleLeu 280  
QY 910 GCACACATGCTGTGTTGGAAGACTGATTGGAAAAGAGCAGAAATTTGAAGAAATT 969  
Db 281 AlaHisAsnGlyLeuValGlyArgLeuIleGlyLysGluGlyArgAsnLeuLysLysIle 300  
QY 970 GAACATGAAAACAGGACCAAGATAACAATCTCATCTTTGAGGATTTGAGCATATACAAC 1029  
Db 301 GluHisGluThrGlyThrLysIleThrIleSerSerLeuGlnAspLeuSerIleTyrAsn 320  
QY 1030 CCGSAAAAGAACCATCACTGTGAAGGGCAGAGTTGAGCGCTGTGCGAGTGTGAGATGAG 1089  
Db 321 ProGluArgThrIleThrValLysGlyThrValGluAlaCysAlaSerAlaGluIleGlu 340  
QY 1090 ATTATGAAGAGCTGCGTGAAGGCTTTGAAAATCATATGCTGCTGTGTACCAACAAGCC 1149  
Db 341 IleMetLysLysLeuArgGluAlaPheGluAsnAspMetLeuAlaValAsn----- 357  
QY 1150 AATCTGTATCCAGGGTTGAACCTCAGCGCACTTGGCATCTTTTCAACAGGACTGTCG 1209  
Db 357 ----- 357  
QY 1210 CTATCTCCACAGAGGGCCCCGGGAGCTCCCCCGCTGCCCGCTTACCACCCCTTCACT 1269  
Db 357 ----- 357  
QY 1270 ACCCACTCCGATACTTCTCCAGCTGTACCCCATCACACAGTTTGGCCGTTCCCGCAT 1329  
Db 358 ThrHisSerGlyTyrPheSerSerLeuTyrProHisGlnPheGlyProPheProHis 377  
QY 1330 CATCACTCTTATCCAGACGAGGAGATTGTGAATCTCTTTCATCCCAACCCAGGCTGTGGC 1389  
Db 378 HisHisSerTyrProGluGlnGluIleValAsnLeuPheIleProThrGlnAlaValGly 397

```
QY 1390 GCCATCATCGGAAGAGGGGCGCACACATCAACAGCTGGCGAGATTCGCGGAGCCTCT 1449
Db 398 AlaIleIleGlyLysGlyAlaHisIleLysGlnLeuAlaArgPheAlaGlyAlaSer 417
QY 1450 ATCAAGATTGCCCCCTGGGAAGGCGCGACAGCTGACGAAAGGATGCTCATCATCCCGG 1509
Db 418 IleLysIleAlaProAlaGluGlyProAspValSerGluArgMetValIleIleThrGly 437
QY 1510 CCACCGAAGCCAGTCAAGGCCCGGAGGACCGATCTTTGGAAACTGAAAGAGAAAC 1569
Db 438 ProProGluAlaGlnPheLysAlaGlnGlyArgIlePheGlyLysLeuLysGluGluAsn 457
QY 1570 TTTCTTAACCCCAAGAGAGTGAAGTGGAGCGCATATCAGAGTGCCTCTTCCACA 1629
Db 458 PhePheAsnProLysGluGluValLysLeuGluAlaHisIleArgValProSerSerThr 477
QY 1630 GCTGCGCGGGTATTGGCAAGAGTGGCAAGACCGTGAACGAACTGCAGAACTTAACCACT 1689
Db 478 AlaGlyArgValIleGlyLysGlyLysThrValAsnGluLeuGlnAsnLeuThrSer 497
QY 1690 GCAGAAGTCATCGTGCCTGTCGACCAAGCCAGATGAAATCAGGAAGTGTATCGTCAGA 1749
Db 498 AlaGluValIleValProArgAspGlnThrProAspGluAsnGluValIleValArg 517
QY 1750 ATTATCGGCACTTCTTCTAGCCAGACTGCACAGCGCAAGATCAGGGAAATTTGACAA 1809
Db 518 IleIleGlyHisPhePheAlaSerGlnThrAlaGlnArgLysIleArgGluIleValGln 537
QY 1810 CAGGTGAAGCAGCAGGAGCAGAAATACCTCAGGAGTGCCTCAGCGCAGCAAG 1866
Db 538 GlnValLysGlnGlnGluGlnLysTyrProGlnGlyValAlaSerGlnArgSerLys 556

RESULT 5
US-10-262-445-40
; Sequence 40, Application US/10262445
; Publication No. US20040014058A1
; GENERAL INFORMATION:
; APPLICANT: Alsebrook II, John
; APPLICANT: Burgess, Catherine
; APPLICANT: Catterton, Elina
; APPLICANT: Chant, John
; APPLICANT: Chaudhuri, Amitabha
; APPLICANT: Edinger, Shlomit
; APPLICANT: Gerlach, Valerie
; APPLICANT: Giot, Loic
; APPLICANT: Gorman, Linda
; APPLICANT: Guo, Xiaojia
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Mezes, Peter
; APPLICANT: Millet, Isabelle
; APPLICANT: Ooi, Chean Eng
; APPLICANT: Patturajan, Meera
; APPLICANT: Rieger, Daniel
; APPLICANT: Spytek, Kimberly
; APPLICANT: Taupier Jr., Raymond J.
; APPLICANT: Zerhusen, Bryan
; APPLICANT: Zhong, Haihong
; APPLICANT: Zhong, Mei
; TITLE OF INVENTION: NOVEL HUMAN PROTEINS, POLYNUCLEOTIDES ENCODING THEM AND METHODS
; FILE OF INVENTION: THE SAME
; FILE REFERENCE: 21402-462D
; CURRENT APPLICATION NUMBER: US/10/262,445
; CURRENT FILING DATE: 2002-10-01
; PRIOR APPLICATION NUMBER: 60/327,454
; PRIOR FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: 60/327,917
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/328,029
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/328,056
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/328,849
; PRIOR FILING DATE: 2001-10-12
```

```
; PRIOR APPLICATION NUMBER: 60/329,414
; PRIOR FILING DATE: 2001-10-15
; PRIOR APPLICATION NUMBER: 60/330,142
; PRIOR FILING DATE: 2001-10-17
; PRIOR APPLICATION NUMBER: 60/341,058
; PRIOR FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: 60/343,629
; PRIOR FILING DATE: 2001-10-24
; PRIOR APPLICATION NUMBER: 60/349,575
; PRIOR FILING DATE: 2001-10-29
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 133
; SOFTWARE: CuraseqList version 0.1
; SEQ ID NO 40
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-262-445-40

Alignment Scores:
Pred. No.: 2,68e-192 Length: 555
Score: 2553.00 Matches: 507
Percent Similarity: 87.81% Conservative: 19
Best Local Similarity: 84.64% Mismatches: 29
Query Match: 41.71% Indels: 44
DB: 15 Gaps: 2

US-09-270-437D-6 (1-3412) x US-10-262-445-40 (1-555)
QY 70 ATGATGAACAAGCTTTTACATCGGAACTGAGCCCGCGCTCACCAGCCGACGACCTCCGG 129
Db 1 MetMetAsnLysLeuPheIleGlyAsnLeuSerProAlaValThrAlaGluAspLeuArg 20
QY 130 CAGCTCTTTGGGGACAGGAAGCTGCCCTGGCGGACAGGTCTCTGCTGAAGTCCGGCTAC 189
Db 21 GlnLeuPheGlyAspArgLysLeuProLeuAlaGlyGlnValLeuLysSerArgTyr 40
QY 190 GCCTTCGTGGACTACCCGACCCAGAACTGGGCCATCCGCCCATCGAGACCCCTCTCGGCT 249
Db 41 AlaPheValAspTyrProAspGlnAsnTrpAlaIleArgThrIleGluThrLeuSerGly 60
QY 250 AAAAGTGAATTGCATGGGAAATCATGGAAGTTGATTACTCAGTCTCTAAAAAGCTAAGG 309
Db 61 GlnValGluLeuHisGlyLysIleMetGluValAspTyrSerValSerIleLysLeuArg 80
QY 310 AGCAGGAAATTCAGATTGCAAAACATCCCTCTCAGTGCAGTGGAGGTGTGGATGGA 369
Db 81 SerArgAsnIleProIleArgAsnIleProProHisLeuGlnTrpGluValLeuAspGly 100
QY 370 CTTTGGCTCAATATGGGACAGTGGAGATGTGGAACTGGAACAAAGTCAACACACACAGAAC 429
Db 101 LeuLeuAlaGlnTyrGlyThrValGluAsnValGluGlnValAsnThrAspThrGluThr 120
QY 430 GCCGTTGTCAACGTCAATATGCAATATGCAACAAGAGAAAGCAAAATAGCCATGGAAGCTA 489
Db 121 AlaValValAsnValThrTyrAlaThrLysGluGluValLysIleAlaMetLysLysLeu 140
QY 490 ACGGGGCATCAGTTTGAGAACTACTCCTTCAGATTTCTACATCCCGATCCGATGAAGGTTG 549
Db 141 SerGlyHisGlnPheGluAsnHisTyrPheLysIleSerTyrIleProAspAspGluVal 160
QY 550 AGCTCCCTTCGCGCCCTCAGCGAGCCCGAGCTGGGACCACTCTTCCCGGAGCAAGGC 609
Db 161 SerCysProSerProProGlnArgAlaGlnArgGlyAspHisSerSerTrpGluGlnGly 180
QY 610 CACGCCCTTGGGGCATTCTCAGGCCAGACAGATTGATTTCCGCTCGGATCCTGGTC 669
Db 181 GlnAlaProGlyGlySerSerGlnAlaArgGlnIleAspPheProLeuArgValLeuPhe 200
QY 670 CCCACCCGTTTGGTGGCCCATCATCGGAAAGGAGGCTTCACCATGAAGACATCACT 729
Db 201 ProThrGlnPheValGlyAlaIleIleGlyLysGluGlyLeuThrIleLysAsnIleThr 220
```



QY 730 AAGCAGACCCAGTCCGGGTAGATATCCATAGAAAGAGAACTCGAGCTCGAGGAG 789  
Db LysGlnSerArgSerArgValAspIleYrArgGlnGluAsnSerArgAlaGluLys 240  
QY 790 CTTGTCAATCCATCCACCCAGAGGGGACTTCTGAAGCATGCCGATGATCTTGAA 849  
Db ProValThrMetHisAlaThrProGluGlyThrSerGluAlaCysArgMetIleLeuGlu 260  
QY 850 ATCATCGAAGAGAGCAGATGAGACCAAACTAGCGGAAGATTCCTCTGAAAATCTTG 909  
Db IleMetGlnLysGluAlaAspGluAlaLysLeuAlaGluGluIleProLeuLysIleLeu 280  
QY 910 GCACAATAGCTGGTGTGAAGACTGATTGAAAAGAGGAGCAAAATTTGAGAAATTT 969  
Db AlaHisAsnGlyLeuValGlyArgLeuIleGlyLysGluGlyArgAsnLeuLysAsn 300  
QY 970 GAACATGAACAGGACCAAGATAACAATCTCATCTTTGCGAGATTGAGCATATACAA 1029  
Db GluHisGluThrGlyThrLysIleThrIleSerSerSerGlnAspLeuSerIleThrAsn 320  
QY 1030 CCGGAAGAACCATCACTGTGAGGCGACAGTTGAGGCTGTGCGAGTGTGAGATAGAG 1089  
Db ProGluArgThrIleThrValLysGlyThrValGluValCysAlaSerAlaGluIleGlu 340  
QY 1090 ATTATGAAGAGCTCGTGGAGGCTTTGAAATGATATGCTGGCTGTTAAACCAAGCC 1149  
Db IleMetLysLysLeuArgGluAlaPheGluAsnAspThrLeuThrValAsn----- 357  
QY 1150 AATCTGATCCAGGGTTGAACCTCAGCGCACTTTGGCATCTTTTCAACAGGACTGTCCGTG 1209  
Db 357 ----- 357  
QY 1210 CTATCTCACCAGAGGCGCGGAGCTCCCGCGTCCCGCTGCGCCCTACACCCCTTCACT 1269  
Db 357 ----- 357  
QY 1270 ACCCACTCCGATATCTTCCAGCGCTGTACCCCATACACCTGTTGGCCGCTCCCGCAT 1329  
Db ThrHisPheGlyTyrPheSerSerLeuTyrProHisArgGlnPheGlyProPheProHis 377  
QY 1330 CATCACTCTTATCCAGAGCAGGAGATTGTGAATCTCTTTCATCCCAACCCAGCGCTGTGGC 1389  
Db HisHisSerTyrProGluGlnGluIleValAsnLeuPheIleProThrGlnGlyValGly 397  
QY 1390 GCGATCATCGGAAGAGGGGCAACATCAAAAGCTGGGAGATTCCCGAGCTCT 1449  
Db AlaIleIleGlyLysLysGlyAlaHisIleLysGlnLeuAlaArgPheValGlyAlaSer 417  
QY 1450 ATCAAGATTGCGCTCGCGAAGGCCAGACGCTCAGCGAAAGGATGCTCATCATCACCGG 1509  
Db IleLysIleAlaProAlaArgSerPro-----LeuArgGlnArgLysValIleIleThrTrp 436  
QY 1510 CCACGGAGCCCGAGTTCAAGCCCGAGGCGGATCTTTGGGAAACTGAAAGAGAAAC 1569  
Db ProProGluSerGlnPheLysAlaGlnGlyArgIlePheGlyLysLeuLysGluLysAsn 456  
QY 1570 TTTCTTAAACCCAAAGAGAGTAGTGAAGTGAAGCGGATATCAGAGTGCCCTCTTCCACA 1629  
Db PhePheAsnProLysGluAspValLysLeuGluThrHisIleArgValProSerSerThr 476  
QY 1630 GCTGCGCGGTGATTGGCAAGGTGGCAAGACCGTGAACGACGACGACGACGACGACG 1689  
Db AlaGlyArgValIleGlyLysGlyLysThrValAsnGluLeuGlnAsnLeuIleSer 496  
QY 1690 GCAGAAGTCACTGCTCGTACCAACAGCCAGATGAAATAGAGGAGTATCGTCCAGA 1749  
Db AlaGluValIleValProArgAspGlnThrProAspGluAsnGluGluMetIleValArg 516  
QY 1750 ATTATCGGCACTCTTCTGCTAGCAGACTCCACAGCGCAAGATCAGGAAATTTGTCAA 1809  
Db IleIleGlyHisPhePheAlaSerGlnThrAlaGlnArgLysIleArgGluIleValGln 536  
QY 1810 CAGGTGAAGCAGCAGGAGCAGAAATACCTCAGGAGTCCGCTCACAGCGCAGCAAG 1866

Db 537 GlnValLysGlnGlnGluGlnLysTyrProGlnGlyValAlaSerGlnArgSerLys 555  
RESULT 6  
US-10-313-986-500  
; Sequence 500, Application US/10313986  
; Publication No. US20030236209A1  
; GENERAL INFORMATION:  
; APPLICANT: Foy, Teresa M.  
; APPLICANT: McNabb, Andria  
; APPLICANT: Watanabe, Yoshihiro  
; APPLICANT: Reed, Steven G.  
; APPLICANT: Wang, Tongtong  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER  
; FILE REFERENCE: 210121.455C19  
; CURRENT APPLICATION NUMBER: US/10/313,986  
; CURRENT FILING DATE: 2002-12-04  
; NUMBER OF SEQ ID NOS: 560  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 500  
; LENGTH: 577  
; TYPE: PXT  
; ORGANISM: Homo sapiens  
US-10-313-986-500  
Alignment Scores:  
Pred. No.: 1,15e-150 Length: 577  
Score: 2025.50 Matches: 396  
Percent Similarity: 78.94% Conservative: 80  
Best Local Similarity: 65.67% Mismatches: 96  
Query Match: 33.09% Indels: 31  
DB: 15 Gaps: 4  
US-09-270-437D-6 (1-3412) x US-10-313-986-500 (1-577)

QY 73 ATGAAACAGCTTTTACATCGGGAACCTGAGCCCCCGCTCACCGCGACGACCTCGCGAG 132  
Db 1 MetAsnLysLeuTyrIleGlyAsnLeuAsnGluSerValThrProAlaAspLeuLys 20  
QY 133 CTCTTTGGGACAGGAAGTCCCTCGCGGAGAGTCTCTGCTGAGTCCGGTACGCC 192  
Db 21 ValPheAlaGluHisLysIleSerTyrSerGlyGlnPheLeuValLysSerGlyTyrAla 40  
QY 193 TTTCTGTGACTACCCCGACCAAGTGGCCATCGGCGCATCGAGACCTCTCGGTA 252  
Db 41 PheValAspCysProAspGluHisTrpAlaMetLysAlaIleGluThrPheSerGlyLys 60  
QY 253 GTGAATTCATGCGGAAATCATGGAAGTTGATTTACTCAGTCTCTAAAAGCTAAGGAGC 312  
Db 61 ValGluLeuGlnGlyLysArgLeuGluIleGluHisSerValProLysLysGlnArgSer 80  
QY 313 AGGAATTCAGATTCGAAATCATCCCTCCTCAGTCTGAGTGGAGTGGTGGTGGTGGTGGT 372  
Db 81 ArgLysIleGlnIleArgAsnIleProGlnLeuArgTrpGluValLeuAspSerLeu 100  
QY 373 TTGCTCAATATGGACAGTGGAGAAATGTGAAACAACTCAACACAGACACAGAAACCGCC 432  
Db 101 LeuAlaGlnTyrGlyThrValGluAsnCysGluGlnValAsnThrGluSerGluThrAla 120  
QY 433 GTTCTCAAGTTCATATGCAACAGAGAAAGAAAGAAATAGCCATGGAGAACTAAGC 492  
Db 121 ValValAsnValThrTyrSerAsnArgGluGlnThrArgGlnAlaIleMetLysLeuAsn 140  
QY 493 GGGCATCAGTTTGGAGTACTCTCTCAAGATTTCTTACATCCCGGATGAAGGTGAGC 552  
Db 141 GlyHisGlnLeuGluAsnHisAlaLeuLysValSerTyrIleProAspGluGlnIleAla 160  
QY 553 TCCCTTCGCGCCCTCAGCAGAGCCCGTGGGACCACTCTTCCCGGAGGAGGCCAC 612  
Db 161 -----GlnGlyProGluAsnGlyArgGlyGlyPheGlySerArgGlyGlnProArg 178  
QY 613 -----GCCCTTGGGGGCACTTCTCAGGCCACAGATTTGATTTCCCGCTG 657

```
Db 179 GlnGlySerProValAlaAlaGlyAlaProAlaLysGlnGlnValAspIleProLeu 198
QY 658 CGGATCTGGTCCCGACCCAGCTTTGGTGGCCATCATCGGAAAGAGGCGCTTGACCAT 717
Db 199 ArgLeuLeuValProThrGlnTrpValGlyAlaIleIleGlyLysGluGlyAlaThrIle 218
QY 718 AAGAATCATCAAGCAGACCCAGTCCCGGTAGATATCCATAGAAAAGAGAACTCTGGA 777
Db 219 ArgAsnIleThrLysGlnThrGlnSerLysIleAspValHisArgLysGluAsnAlaGly 238
QY 778 GCTGAGAGAACCTGTGTCACCATCCATGCCACCCAGAGGGGACTTCTGAACATGCCGC 837
Db 239 AlaAlaGluLysAlaIleSerValHisSerThrProGluGlyCysSerSerAlaCysLys 258
QY 838 ATGATCTTGAATCATGCAAGAAAGAGGAGATAGACCAAACTAGCCGAAGATTCCT 897
Db 259 MetIleLeuGluIleMetHisLysGluAlaLysAspThrLysThrAlaAspGluValPro 278
QY 898 CTGAAATCTTGGCACACAATGCTGTGGTTGGAAGACTGATTGGAAAAGAGCAGAAAT 957
Db 279 LeuLysIleLeuAlaHisAsnAsnPheValGlyArgLeuIleGlyLysGluGlyArgAsn 298
QY 958 TTGAAGAAATTTGAACATGAACAGGGACCAAGATAACAATCTCATTTTGAGAGTTTG 1017
Db 299 LeuLysLysValGluGlnAspThrGluThrLysIleThrIleSerSerLeuGlnAspLeu 318
QY 1018 AGCATATACACCCGAAAGAACCATCACTGTGAGGGGACAGTTGAGGCCCTGTGCCAGT 1077
Db 319 ThrLeuTrpAsnProGluArgThrIleThrValLysGlyAlaIleGluAsnCysArg 338
QY 1078 GCTGACATAGAGATTATGAAGAGCTGCTGAGGCTTTGAAATATGATGCTGGCTGT 1137
Db 339 AlaGluGlnGluIleMetLysLysValArgGluAlaTrpGluAsnAspValAlaAlaMet 358
QY 1138 AACCAACAGCCCAATCTGATCCAGGGTTGAACCTCAGCCGCACTTGGCATTTTCAACA 1197
Db 359 SerLeuGlnSerHisLeuIleProGlyLeuAsnLeuAlaAlaValGlyLeuPheProAla 378
QY 1198 GGAATGTCGTGTATCTCCACAGCAGGGCCCCGGAGCTCCCCCGCTGCCCTAC 1257
Db 379 SerSerSerAlaValProPro-----ProSerSerValThrGlyAlaAlaProTrp 396
QY 1258 CACCCCTTCACTACCCACTCCGGATACTTCTCCAGCCTGTATCCCCCATCACCAGTTTGGC 1317
Db 397 SerSerPheMet----- 400
QY 1318 CGGTTCCCGCATCATCACTTTATCCAGAGCAGAGATTGTGAATCTCTTCATCCCAAC 1377
Db 401 -----GlnAlaProGluGlnGluMetValGlnValPheIleProAla 414
QY 1378 CAGGCTGTGGCGCCATCATCGGGAAGAGGGGCAACATCAACAGCTGGCGAGATTC 1437
Db 415 GlnAlaValGlyAlaIleIleGlyLysGlyGlnHisLysLysGlnLeuSerArgPhe 434
QY 1438 GCCGAGACCTCTCAAGATTGCTCCCTCGGAGAGCCAGACTCAGCGAAAAGAGTGTGC 1497
Db 435 AlaSerAlaSerIleLysIleAlaProProGluThrProAspSerLysValArgMetVal 454
QY 1498 ATCATACCGGGCCACCGGAAGCCAGTTCAAGGCCCAGGACGGATCTTTGGGAAACTG 1557
Db 455 IleIleThrGlyProProGluAlaGlnPheLysAlaGlnGlyArgIleTrpGlyLysLeu 474
QY 1558 AAGAGGAAACTTCTTTTAAACCCAAAGAGAGTGAAGCTGAAGCGCATATCACAGTG 1617
Db 475 LysGluGluAsnPhePheGlyProLysGluGluValLysLeuGluThrHisIleArgVal 494
QY 1618 CCCTCTTCCACAGCTGCGGGGTGATTGGCAAGGTGGCAAGACCGTGAACGAAGTCCAG 1677
Db 495 ProAlaSerAlaAlaGlyArgValIleGlyLysGlyGlyLysThrValAsnGluLeuGln 514
QY 1678 AACTTAACGAGTACAGAGTCACTGCTCGTGACCAACCGCAGTGAATGAGGAA 1737
```

```
Db 515 AsnLeuThrAlaAlaGluValValProArgAspGlnThrProAspGluAsnAspGln 534
QY 1738 GTGATCGTCAGAAATTTATCGGCACCTCTTTGTAGCCAGACTGCACAGCGCAAGATCAG 1797
Db 535 ValIleValLysIleIleGlyHisPheTrpAlaSerGlnMetAlaGlnArgLysIleArg 554
QY 1798 GAAATTTACACAGCTGAAGCAGCAGGACGAGAAATACCTTCAGGGAGTCCCTCACAG 1857
Db 555 AspIleLeuAlaGlnValLysGlnHisGlnLysGlyGlnSerAsnGlnAlaGlnAla 574
QY 1858 CCGACCAAG 1866
Db 575 ArgArgLys 577
RESULT 7
US-09-873-637-2
; Sequence 2, Application US/09873637
; Patent No. US20020061543A1
; GENERAL INFORMATION:
; APPLICANT: Ross, Jeffrey
; TITLE OF INVENTION: THE C-MYC CODING REGION DETERMINANT-BINDING PROTEIN
; TITLE OF INVENTION: (CRD-BP) AND ITS NUCLEIC ACID SEQUENCE
; FILE REFERENCE: 960296.95131
; CURRENT APPLICATION NUMBER: US/09/873,637
; CURRENT FILING DATE: 2001-06-04
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 577
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-873-637-2
Alignment Scores:
Pred. No.: 3,41e-150 Length: 577
Score: 2019.50 Matches: 394
Percent Similarity: 78.94% Conservative: 82
Best Local Similarity: 65.34% Mismatches: 96
Query Match: 32.99% Indels: 31
DB: 9 Gaps: 4
US-09-270-437D-6 (1-3412) x US-09-873-637-2 (1-577)
```

```
QY 73 ATGAACAAGCTTTATACATCGGGAACCTGAGCCCGCGTCACCGCCGACGACCTTCGGGAG 132
Db 1 MetAsnLysLeuTrpIleGlyAsnLeuAsnGluSerValThrProAlaAspLeuGluLys 20
QY 133 CTCCTTTGGGACAGGAAGCTGCCCTGGCGGACAGGTCTCTGTAAGTCCGGCTACGCC 192
Db 21 ValPheAlaGluHisLysIleSerTrpSerGlyGlnPheLeuValLysSerGlyTrpAla 40
QY 193 TTCGTGGACTACCCGACAGAACTGGGCGCATCCGCGCCATCGAGACCTCTCGGGTAAA 252
Db 41 PheValAspCysProAspGluHisTrpAlaMetLysAlaIleGluThrPheSerGlyLys 60
QY 253 GTGGAATTCATGGGAAAATCATGGAAGTTGATTCTACTCTCTATAAAGCTTAAGAGC 312
Db 61 ValGluLeuGlnGlyLysArgLeuGluMetGluHisSerValProLysLysGlnArgSer 80
QY 313 AGGAAATTCAGATTCGAAACATCCCTCTCCTACCTCAGTGGGAGGTGTGGATGACATT 372
Db 81 ArgLysIleGlnIleArgAsnIleProGlnLeuArgTrpGluValLeuAspSerLeu 100
QY 373 TTGGCTCAATATGGGACAGTGGAGATGTGGAACTCAAGTCAACACACACAGAAACCGCC 432
Db 101 LeuAlaGlnTrpGlyThrValGluAsnCysGluGlnValAsnThrGluSerGluThrAla 120
QY 433 GTTGTCAACGTCACATATATGCAACAGAGAAAGCAAAATATGCCATGGAGAAGCTAAGC 492
Db 121 ValValAsnValThrTrpSerAsnArgGluGlnThrAlaIleMetLysLeuAsn 140
QY 493 GGGCATCAGTTTGAAGACTACTCTTCAAGATTTCCTACATCCCGATGAAGAGGTGAGC 552
```

Db 141 GlyHisGlnLeuGluAsnHisAlaLeuLysValSerTyrIleProAspGluGlnIleThr 160  
QY 553 TCCCTTCCGCCCCCTCAGCGACCCAGCGTGGGGACCACTCTCCGGGAGAGAGCCAC 612  
Db 161 -----GlnGlyProGluAsnGlyArgGlyGlyPheGlySerArgGlyGlnProArg 178  
QY 613 -----GCCCTCGGGGGCACTCTCCAGCCAGCAGACAGATTGATTCCCGCTG 657  
Db 179 GlnGlySerProValAlaAlaGlyAlaProAlaLysGlnGlnProValAspIleProLeu 198  
QY 658 CGATCTCGTGGTCCCAACCCAGTTCTGGTGCATCATCGGAAGAGGCGCTTGACCATF 717  
Db 199 ArgLeuLeuValProThrGlnTyrValGlyAlaIleIleGlyLysGluGlyAlaThrIle 218  
QY 718 AAGACATCACTAAGCAGACCCAGTCCCGGTAGATATCCATAGAAAAGAGAACTCTGA 777  
Db 219 ArgAsnIleThrLysGlnThrGlnSerLysIleAspValHisArgLysGluAsnAlaGly 238  
QY 778 GCTCAGAGAAAGCCCTGTCAACATCCATGCCACCCAGAGGGGACTTCTGAAGCATGCCGC 837  
Db 239 AlaAlaGluLysAlaIleSerValHisSerThrProGluGlyCysSerSerAlaCysLys 258  
QY 838 ATGATCTTGAATCATCAGAAAGAGCAGATGAGACCAACTAGCCAGAGAGATTCT 897  
Db 259 MetIleLeuGluIleMetHisLysGluAlaLysAspThrLysThrAlaAspGluValPro 278  
QY 898 CTGAAATCTTGGCACACAATGGCTTGGTGAAGACTGATTGGAAAAGAGGCGAAAT 957  
Db 279 LeuLysIleLeuAlaHisAsnAsnPheValGlyArgLeuIleGlyLysGluGlyArgAsn 298  
QY 958 TTGAAGAAATTTGAATCAACAGGACCAAGATAACAATCTCATCTTTCCAGATTG 1017  
Db 299 LeuLysLysValGluGlnAspThrGluThrLysIleThrIleSerSerLeuGlnAspLeu 318  
QY 1018 AGCATATACACCCGAAAGAACCATCACTGTGAAGGCGACAGTTGAGCCCTGTGCCAGT 1077  
Db 319 ThrLeuTyrAsnProGluArgThrIleThrValLysGlyAlaIleGluAsnCysCysArg 338  
QY 1078 GCTGAGATAGAGATTATCAAGACTGCTGAGGCTTTGAAATGATATGCTGCTGTT 1137  
Db 339 AlaGluGlnGluIleMetLysLysValArgGluAlaTyrGluAsnAspValAlaAlaMet 358  
QY 1138 AACCAACAAGCAATCTCATCCAGGGTGTGAACCTCAGCGACTTGGCATCTTTCAACA 1197  
Db 359 SerLeuGlnSerHisLeuIleProGlyLeuAsnLeuAlaValAlaGlyLeuPheProAla 378  
QY 1198 GGACTGTCCTGCTATCTCCACAGACAGGCGCCCGGAGCTCCCGCTGCCCCCTTAC 1257  
Db 379 SerSerSerAlaValProProPro-----ProSerSerValThrGlyAlaAlaProTyr 396  
QY 1258 CACCCCTTCACTACCCACTCCGGATCTTCTCCAGCCTGTACCCCGCATCACAGTTTGGC 1317  
Db 397 SerSerPheMet----- 400  
QY 1318 CCGTTCGCGCATCATCTTATCCAGAGCAGAGATTGTGAATCTCTTTCATCCCAACC 1377  
Db 401 -----GlnAlaProGluGlnGluMetValGlnValPheIleProAla 414  
QY 1378 CAGCTGTGGCGGCATCATCGGAAGAGGGGCGACATCAACAGCTGGCGAGATTC 1437  
Db 415 GlnAlaValGlyAlaIleIleGlyLysGlyGlnHisIleLysGlnLeuSerArgPhe 434  
QY 1438 GCGGAGCTCTATCAGATTGCCCTCGGAAGGCCAGCAGCTCAGGAAAGATGGTC 1497  
Db 435 AlaSerAlaSerIleLysIleAlaProProGluThrProAspSerLysValArgMetVal 454  
QY 1498 ATCATCACCGGGCCACCGGAAGCCAGTTCAAGGCCAGAGGACGATCTTTGGGAAACTG 1557  
Db 455 ValIleThrGlyProProGluAlaGlnPheLysAlaGlnGlyArgIleTyrGlyLysLeu 474  
QY 1558 AAAGAGGAAACTTCTTTAACCCCAAGAGAGAGTGAAGCTGGAAGCGCATATCAGAGTG 1617  
Db 475 LysGluGluAsnPhePheGlyProLysGluGluValLysLeuGluThrHisIleArgVal 494

QY 1618 CCCTCTTCCACAGCTGGCCGGGTGATTGGCAAAAGTGGCAGACCGTGAACGAACTGCAG 1677  
Db 495 ProAlaSerAlaAlaGlyArgValIleGlyLysGlyGlyLysThrValAsnGluLeuGln 514  
QY 1678 AACTTAACCACTGGCAGAGTCTATCGTCTGTGACCAAAACCCAGATGAAAATGAGNA 1737  
Db 515 AsnLeuThrAlaAlaGluValValProArgAspGlnThrProAspGluAsnAspGln 534  
QY 1738 GTGATCGTCAGAAATATCGGGCAGCTTCTTTGTAGCCAGACTGCACAGCGCAAGATCAGG 1797  
Db 535 ValIleValLysIleIleGlyHisPheTyrAlaSerGlnMetAlaGlnArgLysIleArg 554  
QY 1798 GAAATTGTACCAAGCTGAAGCAGCAGCAGCAAAATACCTCAGGGAGTGCCTTCACAG 1857  
Db 555 AspIleLeuAlaGlnValLysGlnHisGlnLysGlyGlnSerAsnLeuAlaGlnAla 574  
QY 1858 CGCAGCAG 1866  
Db 575 ArgArgLys 577  
RESULT 8  
US-10-313-986-486  
; Sequence 486, Application US/10313986  
; Publication No. US20030236209A1  
; GENERAL INFORMATION:  
; APPLICANT: Foy, Teresa M.  
; APPLICANT: McNabb, Andria  
; APPLICANT: Watanabe, Yoshihiro  
; APPLICANT: Reed, Steven G.  
; APPLICANT: Wang, Tongcong  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER  
; FILE REFERENCE: 210121.455C19  
; CURRENT APPLICATION NUMBER: US/10/313,986  
; CURRENT FILING DATE: 2002-12-04  
; NUMBER OF SEQ ID NOS: 560  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 486  
; LENGTH: 589  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-313-986-486  
Alignment Scores:  
Pred. No.: 3,64e-146 Length: 589  
Score: 1968.50 Matches: 395  
Percent Similarity: 77.69% Conservative: 82  
Best Local Similarity: 64.33% Mismatches: 106  
Query Match: 32.16% Indels: 31  
DB: 15 Gaps: 10  
US-09-270-437D-6 (1-3412) x US-10-313-986-486 (1-589)  
QY 73 ATGACAGCTTTATCATCGGGAACCTGAGCCCCCGCTCACCGCGCAGCCTCCGGCAG 132  
Db 1 MetAsnLysLeuTyrIleGlyAsnLeuSerGluAsnAlaProSerAspLeuGluSer 20  
QY 133 CTCTTTGGGACAGGAAGCTGCCCTGGCGGACAGGTCCTGCTGAAGTCCGGCTACGCC 192  
Db 21 IlePheLysAspAlaLysIleProValSerGlyProPheLeuValLysThrGlyTyrAla 40  
QY 193 TTCGTGGACTACCCCGACACCACTGGGCCCATCGCGCCATCGAGACCTCTCGGGTAAA 252  
Db 41 PheValAspCysProAspGluSerTrpAlaLeuLysAlaIleGluAlaLeuSerGlyLys 60  
QY 253 GTGGAATTCATGGGAAAATCATGGAAGTTGATTCTCAGTCTCTAAAAAGCTAAGAGC 312  
Db 61 IleGluLeuHisGlyLysProIleGluValGluHisSerValProLysArgGlnArgIle 80  
QY 313 AGGAAATTCAGATTGAAACATCCCTCCTCAGCTGAGTGGGAGGTGTGGATGACTT 372  
Db 81 ArgLysLeuGlnIleArgAsnIleProProHisLeuGlnTrpGluValLeuAspSerLeu 100



```

Db      1 MetAsnLysLeuTyrlleGlyAsnLeuSerGluAsnAlaAlaProSerAspLeuGluSer 20
QY      133 CTCITTTGGGACAGGAAGCTGCCCTGGGGGACAGGTCCTGCTGAAGTCGGGTACGCC 192
Db      21 IlePheLysAspAlaLysIleProValSerGlyProPheLeuValLysThrGlyTyrAla 40
QY      193 TCTGGGACTACCCGACAGAACTGGGCCATCCGCGCATCGAGACCTCTCCGGTAAA 252
Db      41 PheValAspCysProAspGluSerIlePheLeuLysAlaIleGluAlaLeuSerGlyLys 60
QY      253 GTGGAATTCATGGGAAATCATGAAGTTGATTACTCAGTCTCTTAAAGAACTAAGAGC 312
Db      61 IleGluLeuHisGlyLysProIleGluValGluHisSerValProLysArgGlnArgIle 80
QY      313 AGGAAATTCAGATTCGAAATCCCTCTCTCCTGAGTGAGGAGTGTTCGATGAGATT 372
Db      81 ArgLysLeuGlnIleArgAsnIleProHisLeuGlnTrpGluValLeuAspSerLeu 100
QY      373 TTGGCTCAATATGGGACAGTGGAGATCTGGAACAAGTCAACACAGACACAGAACCGCC 432
Db      101 LeuValGlnTyrGlyValValGluSerCysGluGlnValAsnThrAspSerGluThrAla 120
QY      433 GTTGTCAACGTACATATCAACAGAGAAGCAAAATAGCCATCGAGAAGCTAAGC 492
Db      121 ValValAsnValThrTyrSerLysAspGlnAlaArgGlnAlaLeuAspLysLeuAsn 140
QY      493 GGCCATCAGTTTGAGAACTACTCCTTCAAGATTTCCTACATCCCGGATGAAGAGTGAGC 552
Db      141 GlyPheGlnLeuGluAsnPheThrLeuLysValAlaTyrIleProAspGluThrAlaAla 160
QY      553 TCCCTCTCCGCCCTCCAG-----CGAGCCAGAGT-----GGGACCACTCTCCCGG 600
Db      161 GlnGlnAsnProLeuGlnGlnProArgGlyArgArgGlyLeuGlyGlnArgGlySerSer 180
QY      601 GAGCAAGCCACGCCCTTGGGGACCTTCTCAGGCCACAGAGATTGATTTCCCGTGGCG 660
Db      181 ArgGlnGly---SerProGlySerValSerLysGlnLysProCysAspLeuProLeuArg 199
QY      661 ATCTGTGTCGCCACCCAGTTTGTGGTCCCATCATCGAAAGAGGGCTTGACCATAAAG 720
Db      200 LeuLeuValProThrGlnPheValGlyAlaIleIleGlyLysGluGlyAlaThrIleArg 219
QY      721 AACATCACTAAGCAGACCCAGTCCCGGTAGATATCCATAGAAAGAGAACTCTCGAGCT 780
Db      220 AsnIleThrLysGlnThrGlnSerLysIleAspValHisArgLysGluAsnAlaGlyAla 239
QY      781 GCAGAGAAGCTGTACCATCCATCCACCCAGAGGGGACTTCTGAAGCATGCGCGATG 840
Db      240 AlaGluLysSerIleThrIleLeuSerThrProGluGlyThrSerAlaAlaCysLysSer 259
QY      841 ATTCTTGAATCATCGCAAGAGGCAGATGAGACCAAACTAGCCGGAAGAGATTCTCTGTG 900
Db      260 IleLeuGluIleMetHisLysGluAlaGlnAspIleLysPheThrGluGluIleProLeu 279
QY      901 AAAATCTTTGGCACACAATGGCTTGGTGGGAAGTGAATGCAAAAGAGGAGGAGAAATTG 960
Db      280 LysIleLeuAlaHisAsnAsnPheValGlyArgLeuIleGlyLysGluGlyArgAsnLeu 299
QY      961 AAGAAATTTGAACATGAACAGGACCAAGATATCAATCTCATCTTTGCGAGTATTCGAGC 1020
Db      300 LysLysIleGluGlnAspThrAspThrLysIleThrIleSerProLeuGlnGluLeuThr 319
QY      1021 ATATACACCCGGAAGAACCATCATCTGTGAAGGCGACAGTTGAGGCTGTGCGCAGTGT 1080
Db      320 LeuTyrAsnProGluArgThrIleThrValLysGlyAsnValGluThrCysAlaLysAla 339
QY      1081 GAGATAGAGATTATGAAGAAGCTCGTGAGGCTTTTGAATATGATATGCTGGCTGTAAAC 1140
Db      340 GluGluGluIleMetLysLysIleArgGluSerTyrGluAsnAspIleAlaSerMetAsn 359
QY      1141 CAAACAGCCAAATCTGATCCCGAGGTGAACCTCAGCGCACTTGGCATCTTTTCAACAGGA 1200

```

```

Db      360 LeuGlnAlaHisLeuIleProGlyLeuAsnLeuAlaLeuGlyLeuPhe----- 376
QY      1201 CTGTCCGTGCTATCTCCACAGCAGGGCCCGCGAGCTCCCGCGCTGCCCTTACCAC 1260
Db      377 -----ProProThrSer-----GlyMetProProProThr----- 386
QY      1261 CCCITCACTACCCACTCCGGATCTTCTCCAGCCTGTACCCCATCACCAGTTTGGCCCG 1320
Db      387 -----SerGlyProProSerAlaMetThrPro-----Pro 396
QY      1321 TTCCCGCATCATCTCTTATCCAGAGCAGAGATTGTGAATCTCTTATCCCAACCCAG 1380
Db      397 TyrProGlnPheGluGln---SerGluThrGlnThrValHisLeuPheIleProAlaLeu 415
QY      1381 GCTGTGGCGCCATCATCGGAAGAGGGGCACACATCAACAGCTGGCAGATTTCGCC 1440
Db      416 SerValGlyAlaIleIleGlyLysGlnGlyGlnHisIleLysGlnLeuSerArgPheAla 435
QY      1441 GGAGCCTCATCAAGATTGCCCTCGCGAAGGCCAGACGTGAGGAAAGGATGTCATC 1500
Db      436 GlyAlaSerIleLysIleAlaProAlaGluAlaProAspAlaLysValArgMetValIle 455
QY      1501 ATCAACCGGGCCACCGAAGCCAGTTCGAAGCCCGAGGACGGATCTTTGGGAAACTGAAA 1560
Db      456 IleThrGlyProProGluAlaGlnPheLysAlaGlnGlyArgIleTyrGlyLysIleLys 475
QY      1561 GAGAAACTTCTTTAAACCCCAAGAGAGTGAAGCTGGAAGCGCATATCAGAGTGGCC 1620
Db      476 GluGluAsnPheValSerProLysGluGluValLysLeuGluAlaHisIleArgValPro 495
QY      1621 TCTTCCACAGTCCCGGTGATTGGCAAGGTGGCAGACCGTGAAGAACTGCAGAAC 1680
Db      496 SerPheAlaAlaGlyArgValIleGlyLysGlyLysThrValAsnGluLeuGlnAsn 515
QY      1681 TTAACCAAGTCAGAAAGTCATCGTCTCGTACCAACAGCCAGATGAAATGAGGAGTG 1740
Db      516 LeuSerSerAlaGluValValProArgAspGlnThrProAspGluAsnAspGlnVal 535
QY      1741 ATCTCAGATTAATTCGGCGCTTCTTTGCTAGCCAGACTGCACAGCGCAGATCAGGAA 1800
Db      536 ValValLysIleThrGlyHisPheTyrAlaCysGlnValAlaGlnArgLysIleGlnGlu 555
QY      1801 ATTCTACAACAGGTGAAGCAGCAGGACAGAAA 1833
Db      556 IleLeuThrGlnValLysGlnHisGlnGlnGln 566

RESULT 10
US-09-850-716A-348
; Sequence 348. Application US/09850716A
; Patent No. US20020115139A1
; GENERAL INFORMATION:
; APPLICANT: Kalos, Michael D.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Retter, Marc W.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.455C15
; CURRENT APPLICATION NUMBER: US/09/850,716A
; NUMBER OF SEQ ID NOS: 440
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 348
; LENGTH: 579
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-850-716A-348

Alignment Scores:
Pred. No.: 2,67e-145 Length: 579
Score: 1957.50 Matches: 387
Percent Similarity: 79.19% Conservative: 81
Best Local Similarity: 65.48% Mismatches: 94
Query Match: 31.98% Indels: 29

```

DB: 9 Gaps: 8  
US-09-270-437D-6 (1-3412) x US-09-850-716A-348 (1-579)  
QY 73 ATGAACAAGCTTTACATCGGGAACCTTGAGCCCGCGCTCACCGCCGACGACCTCCGGCAG 132  
DB 1 MetAsnLysLeuTyrIleGlyAsnLeuSerGluAsnAlaAalaProSerAspLeuGluSer 20  
QY 133 CTCCTTTGGGACAGGAAGCTGCCCTCGGGGACAGGCTCTGCTGAAGTCCGGCTACGGC 192  
DB 21 IlePheLysAspAlaLysIleProValSerGlyProPheLeuValLysThrGlyTyrAla 40  
QY 193 TTCGTGGACTACCCGACAGAACTGGGCCATCCGCGCCATCGAGACCTCTCGGTA 252  
DB 41 PheValAspCysProAspGluSerTrpAlaLeuLysAlaIleGluAlaLeuSerGlyLys 60  
QY 253 GTGGATTGCTATGGGAAATCATGGAAGTTGATTCTCATGCTCTCTAAAAGTAAGGAGC 312  
DB 61 IleGluLeuHisGlyLysProIleGluValGluHisSerValProLysArgGlnArgIle 80  
QY 313 AGGAAATTCAGATTCGAAACATCCCTCTCACTGCGAGTGGAGGTGTGGATGACTT 372  
DB 81 ArgLysLeuGlnIleArgAsnIleProProHisLeuGlnTrpGluValLeuAspSerLeu 100  
QY 373 TTGGCTCAATATGGGACAGTGGAGATGTGGAACAAGTCAACACAGACACAGAAACCGC 432  
DB 101 LeuValGlnTyr-GlyValValGluSerCysGluGlnValAsnThrAspSerGluThrAla 120  
QY 433 GTTGTCAACGTCACATATCAACAAGAGAAAGCAAAAATAGCCATGGAGAGCTAAGC 492  
DB 121 ValValAsnValThrTyrSerSerLysAspGlnAlaArgGlnAlaLeuAspLysLeuAsn 140  
QY 493 GGGCATCATGTTGAGAACTACTCTTCAGATTCTCATATCCGATCCGATCAAGAGGTGAGC 552  
DB 141 GlyPheGlnLeuGluAsnPheThrLeuLysValAlaTyrIleProAspGluThrAlaAla 160  
QY 553 TCCCTCTCCCGCTCAG-----CGAGCCAGCGT-----GGGACCACTCTTCCCGG 600  
DB 161 GlnGlnAsnProLeuGlnGlnProArgGlyArgArgGlyLeuGlyGlnArgGlySerSer 180  
QY 601 GAGCAAGGCCACCGCCCTCGGGGCACTTCTCAGGCCAGACAGATTGATTTCCGCTGGG 660  
DB 181 ArgGlnGly---SerProGlySerValSerLysGlnLysProCysAspLeuProLeuArg 199  
QY 661 ATCTCGTCCACCCAGTTCTGTGGTCCATCCTCGAAAGAGAGGCTTGACCTAAG 720  
DB 200 LeuLeuValProThrGlnPheValGlyAlaIleIleGlyLysGluGlyAlaThrIleArg 219  
QY 721 AACATCACTAAGCAGACCCCGGTAGATATCCATAGAAAAGAGAACTCTGGAGCT 780  
DB 220 AsnIleThrLysGlnThrGlnSerLysIleAspValHisArgLysGluAsnAlaGlyAla 239  
QY 781 GCAGAGAAGCTGTACCATCCATGCCACCCGAGGGGACTTCTGAAGATGCCCGCATG 840  
DB 240 AlaGluLysSerIleThrIleLeuSerThrProGluGlyThrSerAlaAlaCysLysSer 259  
QY 841 ATTCTTGAATCATGCAGAAAGAGGAGATGAGACCAAACTAGCCGACAGATTCTCTCG 900  
DB 260 IleLeuGluIleMetHisLysGluAlaGlnAspIleLysPheThrGluGluIleProLeu 279  
QY 901 AAAATCTTGGCCACCAATGGCTTGGTTGGAAGACTGATTGGAAGAGAGCGAGAAATTTG 960  
DB 280 LysIleLeuAlaHisAsnAsnPheValGlyArgLeuIleGlyLysGluGlyArgAsnLeu 299  
QY 961 AGAAAATGAACATGAACAGGACCAAGATTAACAATCTCATCTTTGAGGATTTGAGC 1020  
DB 300 LysLysIleGluGlnAspThrAspThrLysIleThrIleSerProLeuGlnGluLeuThr 319  
QY 1021 ATATACACCCGGAAGAACCATCACTGTGAAGGACACATGTTGAGCGCTGTGCGAGTCT 1080  
DB 320 LeuTyrAsnProGluArgThrIleThrValLysGlyAsnValGluThrCysAlaLysAla 339  
QY 1081 GAGATAGAGATTATGAAGAAGCTGCGTAGGCGCTTTGAAAATGATATGCTGGCTGTTAAC 1140

DB 340 GluGluGluIleMetLysLysIleArgGluSerTyrGluAsnAspIleAlaSerMetAsn 359  
QY 1141 CAACAAGCAATCTCATCCAGGGTTGAACCTCAGCGACCTTGGCATCTTTTCAACAGA 1200  
DB 360 LeuGlnAlaHisIleIleProGlyLeuAsnLeuAsnAlaLeuGlyLeuPhe----- 376  
QY 1201 CTGTCTGGTCTATCTCCACACAGAGGGGCCCCCGGAGCTCCCCCGCTGCCCTTACCAC 1260  
DB 377 -----ProProThrSer-----GlyMetProProProThr----- 386  
QY 1261 CCCTTCACTACCCACTCCGGTACTTCTCCAGCTTACCCCTTACCCCATCACAGTTTGGCCGC 1320  
DB 387 -----SerGlyProProSerAlaMetThrPro-----Pro 396  
QY 1321 TTCGCGCATCATCTCTTATCCAGAGCAGAGATTGGAATCTCTCATCCCAACCCAG 1380  
DB 397 TyrProGlnPheGluGln---SerGluThrGluThrValHisLeuPheIleProAlaLeu 415  
QY 1381 GCTGTGGCGCCCATCATCGGGAAGAGGGGCACACATCAACAGCTGCGGAGATTCGCC 1440  
DB 416 SerValGlyAlaIleIleGlyLysGlnGlyHisIleLysGlnLeuSerArgPheAla 435  
QY 1441 CGAGCTCTTATCAAGATTGCCCTCGGGAAGGCCAGAGCTCAGCGAAAGGATGCTCATC 1500  
DB 436 GlyAlaSerIleLysIleAlaProAlaGluAlaProAspAlaLysValArgMetValIle 455  
QY 1501 ATCAGCGGCCACCGAGGCCAGTTCAGGCGCCAGGACGATCTTTGGGAAACTGAAA 1560  
DB 456 IleThrGlyProProGluAlaGlnPheLysAlaGlnGlyArgIleTyrGlyLysIleLys 475  
QY 1561 GAGGAAAACCTCTTTAAACCCCAAGAAAGTGAAGCTGGAAGCGCATATCAGAGTGCC 1620  
DB 476 GluGluAsnPheValSerProLysGluGluValLysLeuGluAlaHisIleArgValPro 495  
QY 1621 TCTTCCACAGCTGGCGGCTGATTGCGCAAGTGCAAGACCGTGAACCAACTGCGAAGC 1680  
DB 496 SerPheAlaAlaGlyArgValIleGlyLysGlyLysThrValAsnGluLeuGlnAsn 515  
QY 1681 TTAAACAGTGCAGAGTCATCTGCTGCTGCTGACCAACCCAGATCAAAATGAGGAAGTG 1740  
DB 516 LeuSerSerAlaGluValValProArgAspGlnThrProAspGluAsnAspGlnVal 535  
QY 1741 ATCTGTCAAGATTATCGGCACTTCTTGTAGCAGACTGCACAGCGCAAGATCAGGGAA 1800  
DB 536 ValValLysIleThrGlyHisPheTyrAlaCysGlnValAlaGlnArgLysIleGlnGlu 555  
QY 1801 ATTGTACACAGGTGAAGCAGCAGGAGCAGAAA 1833  
DB 556 IleLeuThrGlnValLysGlnHisGlnGlnGln 566

RESULT 11

US-09-897-778-348  
; Sequence 348, Application US/09897778  
; Patent No. US20020147143A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Tongtong  
; APPLICANT: Marnerakis, Margarita  
; APPLICANT: Fanger, Gary R.  
; APPLICANT: Vedwick, Thomas S.  
; APPLICANT: Carter, Darrick  
; APPLICANT: Watanabe, Yoshihiro  
; APPLICANT: Henderson, Robert A.  
; APPLICANT: Pechtham, David W.  
; APPLICANT: Fanger, Neil  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER  
; FILE REFERENCE: 210121.455C16  
; CURRENT APPLICATION NUMBER: US/09/897,778  
; CURRENT FILING DATE: 2001-06-28  
; NUMBER OF SEQ ID NOS: 467  
; SOFTWARE: FASTSEQ for Windows Version 4.0  
; SEQ ID NO 348

```
; LENGTH: 579
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-897-778-348

Alignment Scores:
Pred. No.: 2,67e-145 Length: 579
Score: 1957,50 Matches: 387
Percent Similarity: 79.19% Conservative: 81
Best Local Similarity: 65.48% Mismatches: 94
Query Match: 31.98% Indels: 29
DB: 9 Gaps: 8

US-09-270-437D-6 (1-3412) x US-09-897-778-348 (1-579)
QY 73 ATGAACAAAGCTTACATCGGAACCTGAGCCCGCCGCTACCGCGCAGACCTCCGGCAG 132
Db 1 MetAsnLysLeuTyrIleGlyAsnLeuSerGluAsnAlaAlaProSerAspLeuGluSer 20
QY 133 CTCCTTGGGGACAGGAAGCTCCCTCGCGGGACAGGTCCTGCTGAAGTCGGCTACGCC 192
Db 21 IlePheLysAspAlaLysIleProValSerGlyProPheLeuValLysThrGlyTyrAla 40
QY 193 TTCTGTGACTACCCCGACAGAACTGGGCCATCGCGCCATCGAGACCTCTCGGGTAAA 252
Db 41 PheValAspCysProAspGluSerTrpAlaLeuLysAlaIleGluAlaLeuSerGlyLys 60
QY 253 GTGGAAATGCGATGGGAAATCATGGAAGTTGATTAAGTCTCTAAAGAGCTAAGGAGC 312
Db 61 IleGluLeuHisGlyLysProIleGluValGluHisSerValProLysArgGlnArgIle 80
QY 313 AGGAAATTCAGATTCGAACATCCCTCCTCACCTGCACTGGAGGCTTGGATGGATT 372
Db 81 ArgLysLeuGlnIleArgAsnIleProProHisLeuGlnTrpGluValLeuAspSerLeu 100
QY 373 TTGGCTCAATATGGGACAGTGGAGAAGTGGAAACAAGTCAACACAGACACAGAAACGCC 432
Db 101 LeuValGlnTyrGlyValValGluSerCysGluGlnValAsnThrAspSerGluThrAla 120
QY 433 GTTGTCAAGTCACATATGCAACAGAGAGAGAGCAAAATATAGCCATGGAGAAGCTAAGC 492
Db 121 ValValAsnValThrTyrSerLysAspGlnAlaArgGlnAlaLeuAspLysLeuAsn 140
QY 493 GGGCATCAGTTTGAGAACTACTCCTTCAGATTTCCTACATCCCGATGAGAGAGTGAGC 552
Db 141 GlyPheGlnLeuGluAsnPheThrLeuLysValAlaTyrIleProAspGluThrAlaAla 160
QY 553 TCCCTCTCGCCCTCCAG-----CGAGCCAGCGT-----GGGGACCACTCTCCCGG 600
Db 161 GlnGlnAsnProLeuGlnProArgGlyArgArgGlyLeuGlyGlnArgGlySerSer 180
QY 601 GAGCAAGCCACGCCCTCGGGGCACCTTCTCAGCCAGACAGATGATTTCCCGCTCGG 660
Db 181 ArgGlnGly---SerProGlySerValSerLysGlnLysProCysAspLeuProLeuArg 199
QY 661 ATCTGTGTCCCCACCCAGCTTGGTGGCCATCATCGGAAGGAGGGCTTACCATTAAG 720
Db 200 LeuLeuValProThrGlnPheValGlyAlaIleIleGlyLysGluGlyAlaThrIleArg 219
QY 721 AACATCACTAAGCAGACCCAGTCCCGGTGATATCCATAGAAAAGAGAACTCTGGAGCT 780
Db 220 AsnIleThrLysGlnThrGlnSerLysIleAspValHisArgLysGluAsnAlaGlyAla 239
QY 781 GCAGAGAAGCTGTCAACATCATGCCACCCAGAGGGGACTTCTGGAAGCATGCCGATG 840
Db 240 AlaGluLysSerIleThrIleLeuSerThrProGluGlyThrSerAlaAlaCysLysSer 259
QY 841 ATTCTTGAATCATGCAGAAAGAGCAGATGAGACCAACTAGCCGAGAGATTCCTCTG 900
Db 260 IleLeuGluIleMetHisLysGluAlaGlnAspIleLysPheThrGluGluIleProLeu 279
QY 901 AAAATCTTGGCACACAATGGCTGTGGTGGAAAGCTGATTGGAAGAGAGGAGCAAAATTG 960
Db 960
```

```
Db 280 LysIleLeuAlaHisAsnAsnPheValGlyArgIleLeuGlyLysGluGlyArgAsnLeu 299
QY 961 AAGAAAATTGAACATGAACAGGACCAAGATAACAATCTCTATCTTTTCAGGATTGAGC 1020
Db 300 LysLysIleGluGlnAspThrAspThrLysIleThrIleSerProLeuGlnGluLeuThr 319
QY 1021 ATATACAAACCCGAAAGAAACCAATCAGTGTGAAGGGCACAGTTGAGCCCTGTGCCAGTCT 1080
Db 320 LeuTyrAsnProGluArgThrIleThrValLysGlyAsnValGluThrCysAlaLysAla 339
QY 1081 GAGATAGAGATTATGAAGAAGCTCGTGAGGCCCTTTGAAAATGATATATCTGCTGTGTTAAC 1140
Db 340 GluGluGluIleMetLysLysIleArgGluSerTyrGluAsnAspIleAlaSerMetAsn 359
QY 1141 CAACAAGCCAACTCTGATCCCAAGGTTGAACCTCAGCGCACTTGGCATCTTTTCAACAGGA 1200
Db 360 LeuGlnAlaHisLeuIleProGlyLeuAsnLeuAlaLeuGlyLeuPhe----- 376
QY 1201 CTGTCCGTGTATCTCCACAGAGAGGCCCGCGGAGCTCCCGCGTGCCTCATCCAC 1260
Db 377 -----ProProThrSer-----GlyMetProProProThr----- 386
QY 1261 CCTTCACTACCCACTCCGATCTCTCCAGCCTGTACCCCATCACCATGTTGGCCCG 1320
Db 387 -----SerGlyProProSerAlaMetThrPro-----Pro 396
QY 1321 TTCCCGCATCATCTCTTATCCAGAGCAGGAGATTGTGAATCTTCTATCTCCCAACCCAG 1380
Db 397 TyrProGlnPheGluGln---SerGluThrGluThrValHisLeuPheIleProAlaLeu 415
QY 1381 GCTGTGGCGGCATCATCTCGGGAAGAGAGGGGCACACATCAACAGCTGGCAGATTGCC 1440
Db 416 SerValGlyAlaIleIleGlyLysGlnGlyHisIleLysGlnLeuSerArgPheAla 435
QY 1441 GGAGCCTCTATCAAGATTGCCCTCGGAGAGGCCAGCGTCAGCGAAAGGATGTCATC 1500
Db 436 GlyAlaSerIleLysIleAlaProAlaGluAlaProAspAlaLysValArgMetValIle 455
QY 1501 ATCAGCGGGCCACCGGAGGCCAGTTTCAAGSCCGGAGCGGATCTTTGGGAAACTGAAA 1560
Db 456 IleThrGlyProProGluAlaGlnPheLysAlaGlnGlyArgIleTyrGlyLysLys 475
QY 1561 GAGGAAAACCTCTTTAACCCCAAGAGAGAGTGAAGCTGGAGCGCATATCAGAGTGC 1620
Db 476 GluGlnAsnPheValSerProLysGluGluValLysLeuGluAlaHisIleArgValPro 495
QY 1621 TCTTCCACAGCTGCCCGGTGATTGGCAAGGTGGCAAGACCGTGAACGAACCTGCAGAAC 1680
Db 496 SerPheAlaAlaGlyArgValIleGlyLysGlyLysThrValAsnGluLeuGlnAsn 515
QY 1681 TTAACCGAGTCAGAAATCATCTGCTCGTGACCAAAACGCCAGATGAAAATGAGGAAGTG 1740
Db 516 LeuSerSerAlaGluValValProArgAspGlnThrProAspGluAsnAspGlnVal 535
QY 1741 ATCTGTGAGATTATCGGGCAGCTCTTCTGCTAGCCAGACTGCACAGCGCAGATCAGGAA 1800
Db 536 ValValLysIleThrGlyHisPheTyrAlaCysGlnValAlaGlnArgLysIleGlnGlu 555
QY 1801 ATTGTACAACAGGTGAAGCAGCAGCAGCAAGAA 1833
Db 556 IleLeuThrGlnValLysGlnHisGlnGlnGln 566
```

## RESULT 12

```
US-09-897-778-446
; Sequence 446, Application US/09897778
; Patent No. US20020147143A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Mainerakis, Margarita
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darriek
; APPLICANT: Watanabe, Yoshihiro
```



; APPLICANT: Henderson, Robert A.  
 ; APPLICANT: Peckham, David W.  
 ; APPLICANT: Fanger, Neil  
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
 ; FILE REFERENCE: 210121.455C16  
 ; CURRENT APPLICATION NUMBER: US/09/897,778  
 ; CURRENT FILING DATE: 2001-06-28  
 ; NUMBER OF SEQ ID NOS: 467  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 446  
 ; LENGTH: 579  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; US-09-897-778-446

Alignment Scores:  
 Pred. No.: 2,67e-145 Length: 579  
 Score: 1957.50 Matches: 387  
 Percent Similarity: 79.19% Conservative: 81  
 Best Local Similarity: 65.48% Mismatches: 94  
 Query Match: 31.98% Indels: 29  
 DB: 9 Gaps: 8

US-09-270-437D-6 (1-3412) x US-09-897-778-446 (1-579)

QY 73 ATGAACAAGCTTTACATCGGACCTGAGCCCGCGCTCACCGCCAGACCTCCGGCAG 132  
 DB 1 MetAsnLysLeuTyrlleGlyAsnLeuSerGluAsnAlaProSerAspLeuGluSer 20  
 QY 133 CTCCTTTGGGACAGGAAGTGCCTCGCGGACAGCTCTGCTGAAGTCCCGCTACGCC 192  
 DB 21 llePheLysAspAlaLysileProValSerGlyProPheLeuVallysthrGlyTyra 40  
 QY 193 TTCGTGGACTACCCGACACAGAACTGGGCCATCCGCGCATCGAGACCTCTCGGTAA 252  
 DB 41 PheValAspCysProAspGluSerTrpAlaLeuLysAlaileGluAlaLeuSerGlyLys 60  
 QY 253 GTGGAAATTCATGCGAAATCATGGAAGTTGATTACTCAGTCTCTAAAAGCTAGGAC 312  
 DB 61 lleGluLeuHisGlyLysProileGluValGluHisSerValProLysArgGlnArgile 80  
 QY 313 AGGAAATTCAGATTGCAACATCCCTCTCACCTGAGTGGGAGGTGTGGATGACTT 372  
 DB 81 ArgLysLeuGlnleArgAsnileProProHisLeuGlnTrpGluValLeuaspSerLeu 100  
 QY 373 TTGGCTCATATGGACAGTGGAGAAATGGAACAGTCAACACAGACACAGAACCCGCC 432  
 DB 101 LeuValGlnTyrglyValValGluSerCysGluGlnValAsnThrAspSerGluThrAla 120  
 QY 433 GTTGTCAAGTCCACATATCAACAGAGAAGCAAAATAGCCATGGAGAAGCTTAAGC 492  
 DB 121 ValValAsnValThrTyrserserLysAspGlnAlaargGlnAlaLeuaspLysLeuAsn 140  
 QY 493 GGGCATCAGTTTGAAGACTACCTTCAAGATTCTCATATCCCGATGAAGAGGTGAGC 552  
 DB 141 GlyPheGlnLeuGluAsnPheThrLeuLysValAlaTyrlleProAspGluThrAlaAla 160  
 QY 553 TCCCTTCCGCCCTCAG-----CGAGCCAGCGT-----GGGACCACTCTCCCGG 600  
 DB 161 GlnGlnAsnProLeuGlnGlnProArgGlyArgGlyGlyGlyGlyGlyGlyGlySer 180  
 QY 601 GAGCAAGGCCAGCCCTCGGGGCATCTCTCAGGCCAGCAGATTGATTCCCGCTCGCG 660  
 DB 181 ArgGlnGly---SerProGlySerValSerLysGlnLysProCysAspLeuProLeuArg 199  
 QY 661 ATCTGGTCCCAACCCAGTTTGTGTGCCATCATCGGAAAGAGGGCTTGACCAATAAG 720  
 DB 200 LeuLeuValProThrGlnPheValGlyAlaileGlyLysGluGlyAlaThrileArg 219  
 QY 721 AACATCACTAAGCAGACCCAGTCCCGGTAGATATCCATAGAAAAGAGAACTCTCGAGCT 780  
 DB 220 AsnilleThrLysGlnThrGlnSerLysileAspValHisArgLysGluAsnAlaGlyAla 239

QY 781 GCAGAGAAGCCCTGTCCACATCCATGCCACCCCGAGAGGAGCTTCTCAACATGCCCATG 840  
 DB 240 AlaGluLysSerileThrileLeuSerThrProGluGlyThrSerAlaAlaCysLysSer 259  
 QY 841 ATTCTTGAATCATCGAGAAAGAGCAGATGATGAGCAACAACTAGCCGCAAGAGATTCTCTG 900  
 DB 260 lleLeuGluileMetHisLysGluAlaGlnAspilleLysPheThrGluGluileProLeu 279  
 QY 901 AAAATCTTGGCACACAATGGCTTGTGAAGACTGATTGGAAGAAAGAGCGAGAATTG 960  
 DB 280 LysilleLeuAlaHisAsnAsnPheValGlyArgLeuileGlyLysGluGlyArgAsnLeu 299  
 QY 961 AAGAAATGTACATGAACAGGACAGAGTGAACATCTCATCTCTTTCAGGATTGAGC 1020  
 DB 300 LysLysilleGluGlnAspThrAspThrLysilleThrileSerProleuGlnGluLeuThr 319  
 QY 1021 ATATACAAACCCGAAAGAACCATCACTGTGAAGGACAGAGTGTGAGCCCTGTCCAGTGT 1080  
 DB 320 LeuTyAsnProGluArgThrileThrValLysGlyAsnValGluThrCysAlaLysAla 339  
 QY 1081 GAGATAGAGATTATGAAGAGCTGCTGAGGCTTTGAAATGATATGCTGCTGTTAATC 1140  
 DB 340 GluGluGluileMetLysileArgGluSerTyrgluAsnAspilleAlaSerMetAsn 359  
 QY 1141 CAACAAGCCAAATCTGATCCAGGGTTGAACCTCAGCGACTTGGCATCTTTTCAACAGGA 1200  
 DB 360 LeuGlnAlaHisLeuileProGlyLeuAsnLeuAsnAlaLeuGlyLysPhe----- 376  
 QY 1201 CTGTCCTGTCTATCTCCACAGAGCGGCCCGCGGAGCTCCCGCGCTCCCGCTACCA 1260  
 DB 377 -----ProProThrSer-----GlyMetProProThr----- 386  
 QY 1261 CCTTCACTACCACTCCGATACCTTCTCCAGCTGTACCCCATCACAGTTTGGCCCG 1320  
 DB 387 -----SerGlyProProSerAlaMetThrPro-----Pro 396  
 QY 1321 TTCCCGCATCATCACTTATCCAGCAGGAGATTGGAATCTCTTTCATCCCAACCCAG 1380  
 DB 397 TyrProGlnPheGluGln---SerGluThrGluThrValHisLeuPheilleProAlaLeu 415  
 QY 1381 GCTGTGGCGCCATCATCGGAAAGAGGGGGCACACATCAACAGCTGGCGAGATTCGCC 1440  
 DB 416 SerValGlyAlaileileGlyLysGlnGlyGlnHisileLysGlnLeuSerArgPheAla 435  
 QY 1441 GGAGCTCTATCAAGATTGCCCTCGGAGCCCGCAGAGCTCAGCGAAAGGATGTCATC 1500  
 DB 436 GlyAlaSerileLysileAlaProAlaGluAlaProAspAlaLysValArgMetValile 455  
 QY 1501 ATCAGCGGCCACCGGAAGCCAGTTCAGAGGCCAGGACGGATCTTTGGGAAACTGAAA 1560  
 DB 456 lleThrGlyProProGluAlaGlnPheLysAlaGlnGlyArgilleTyrglyLysileLys 475  
 QY 1561 GAGGAAACTTCTTTAACCCCAAGAAAGTGAAGCTGGAAGCGCATATCAGAGTGCC 1620  
 DB 476 GluGluAsnPheValserProLysGluGluValLysLeuGluAlaHisileArgValPro 495  
 QY 1621 TCTTCCAGCTGGCGGGTGTGTTGGAAAGTGGCAGACCGCTGAACGAACTGCAGAAC 1680  
 DB 496 SerPheAlaAlaGlyArgValileGlyLysGlyLysThrValAsnGluLeuGlnAsn 515  
 QY 1681 TTAAACAGTGCAGAAAGTTCATCGTCTGTGACCAAAAGCCAGATGAAATCAGGAAGTG 1740  
 DB 516 LeuSerSerAlaGluValValProArgAspGlnThrProAspGluAsnAspGlnVal 535  
 QY 1741 ATCTGCAGAAATTATCGGCACTTCTTGTGTCACCACTGCAGCGCAAGATCAGGAA 1800  
 DB 536 ValValLysileThrGlyHisPheTyraLysGlnValAlaGlnArgLysileGlnGlu 555  
 QY 1801 ATTGTACAACAGGTGAAGCAGCAGCAGCAGAAA 1833  
 DB 556 lleLeuThrGlnValLysGlnHisGlnGlnGln 566



```
RESULT 13
US-09-897-778-449
; Sequence 449, Application US/09897778
; Patent No. US20020147143A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Marnerakis, Margarita
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedwick, Thomas S.
; APPLICANT: Carter, Darriack
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Henderson, Robert A.
; APPLICANT: Peckham, David W.
; APPLICANT: Fanger, Neil
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C16
; CURRENT APPLICATION NUMBER: US/09/897,778
; CURRENT FILING DATE: 2001-06-28
; NUMBER OF SEQ ID NOS: 467
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 449
; LENGTH: 579
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-897-778-449

Alignment Scores:
Pred. No.: 2,67e-145 Length: 579
Score: 1957.50 Matches: 387
Percent Similarity: 79.19% Conservative: 81
Best Local Similarity: 65.48% Mismatches: 94
Query Match: 31.98% Indels: 29
DB: 9 Gaps: 8

US-09-270-437D-6 (1-3412) x US-09-897-778-449 (1-579)
QY 73 ATGAACAGCTTTACATCGGAACCTGAGCCCGCCCGTCCAGCGCAGACCTCCGGCAG 132
Db 1 MetAsnLysLeuTyrlleGlyAsnLeuSerGluAsnAlaLaProSerAspLeuSer 20
QY 133 CTCTTTGGGGACAGGAAGCTGCCCTGGCGGACAGAGTCTGCTGAAGTCCGGGTACGCC 192
Db 21 IlePheLysAspAlaLysIleProValSerGlyProPheLeuValLysThrGlyTyrla 40
QY 193 TTGCTGAGTACCCCGACAGACTGGGCCATCCGCGCATCGAGACCCCTTCGGGTAAA 252
Db 41 PheValAspCysProAspGluSerTrpAlaLeuLysAlaIleGluAlaLeuSerGlyLys 60
QY 253 GTGGAATTGCATGGGAAATCATGAAGTTGATTTACTAGTCTCTAAAGCTAAGGAGC 312
Db 61 IleGluLeuHisGlyLysProIleGluValGluHisSerValProLysArgGlnArgIle 80
QY 313 AGGAAATTCAGATTCTGAACATCCCTCTCACCTGAGTGGGAGGTCTTGATGGAGCTT 372
Db 81 ArgLysLeuGlnIleArgAsnIleProProHisLeuGlnTrpGluValLeuAspSerLeu 100
QY 373 TTGCTCMAATGGACAGTGGAGATGTGGAACAGTCAACAGACACAGACAGAACCGCC 432
Db 101 LeuValGlnTyrlleGlyValValGluSerCysGluGlnValAsnThrAspSerGluThrAla 120
QY 433 GTTGTCAACGTCACATATGCAACAGAGCAAGCAAAATAGCCATCGAGAGCTAAGC 492
Db 121 ValValAsnValThrTyrlSerSerLysAspGlnAlaArgGlnAlaLeuAspLysLeuAsn 140
QY 493 GGGCATCAGTTTGAGAACTACTCTTCAAGATTTCTCATCTCCCGGATGAAGAGGTGAGC 552
Db 141 GlyPheGlnLeuGluAsnPheThrLeuLysValAlaTyrlleProAspGluThrAla 160
QY 553 TCCCTTCGCCCTCTCAG-----CGAGCCCGAGGT-----GGGACCATCTTCCCGG 600
Db 161 GlnGlnAsnProLeuGlnGlnProArgGlyArgGlyLeuGlyGlnArgGlySerSer 180
```

```
QY 601 GAGCAAGSCCAGCCCTCGGGGCACTTCTCAGGCCACAGATTGATTTCCCGCTGGCG 660
Db 181 ArgGlnGly---SerProGlySerValSerLysGlnLysProCysAspLeuProLeuArg 199
QY 661 ATCTGTGTCGCCACCCAGTTTGTGGTCCCATCATCGGAAGAGGGGCTTGACCAATAAG 720
Db 200 LeuLeuValProThrGlnPheValGlyAlaIleleGlyLysGluGlyValaThrIleArg 219
QY 721 AACATCACTAAGCAGACCCAGTCCCGGGTAGATATCCATAGAAAAGAGAACTCTGGAGCT 780
Db 220 AsnIleThrLysGlnThrGlnSerLysIleAspValHisArgLysGluAsnAlaGlyAla 239
QY 781 CGAGAGAAGCTGTCACCATCCATCCACCCAGAGGGGACTTTCGAAGCATGCGCGCATG 840
Db 240 AlaGluLysSerIleThrIleLeuSerThrProGluGlyThrSerAlaAlaCysLysSer 259
QY 841 ATTCTTGAAATCATGTCAGAAAGAGCGAGATGAGCCAAACTAGCCGAAGAGATTCCTCTG 900
Db 260 IleLeuGluIleMetHisLysGluAlaGlnAspIleLysPheThrGluGluIleProLeu 279
QY 901 AAAATCTTGGCACACAATGGCTTGGTTGGAAGACTGATTGGAAAAGAGAGGAGAAATTTG 960
Db 280 LysIleLeuAlaHisAsnAsnPheValGlyArgLeuIleGlyLysGluGlyArgAsnLeu 299
QY 961 AAGAAATTTGAACATGAAACAGGACCAAGATACAAATCTCATCTTTTCAGAGATTTGAGC 1020
Db 300 LysLysIleGluGlnAspThrAspThrLysIleThrIleSerProLeuGlnGluLeuThr 319
QY 1021 ATATACAAACCCGGAAGAAACCATCATCTGTAAGGCGCACAGTTGAGGCTGTGCCAGTCT 1080
Db 320 LeuTyrlAsnProGluArgThrIleThrValLysGlyAsnValGluThrCysAlaLysAla 339
QY 1081 GAGATAGAGATTATGAAGAAGCTCGTGAAGCCCTTTGAAAATGATATGCTGGCTGTTAAC 1140
Db 340 GluGluGluIleMetLysLysIleArgGluSerTyrlGluAsnAspIleAlaSerMetAsn 359
QY 1141 CAACAGCAATCTGATCCCGAGGCTTGACCTCAGCGCAGTGGCATCTTTTCAACAGGA 1200
Db 360 LeuGlnAlaHisLeuIleProGlyLeuAsnLeuAsnAlaLeuGlyLeuPhe----- 376
QY 1201 CTGTCCGTGCTATCTCCACACAGAGGCGCCCGCGAGCTCCCGCGCTGCCCTTACCAC 1260
Db 377 -----ProProThrSer-----GlyMetProProProThr----- 386
QY 1261 CCCTTCACTACCCACTCCGATACTTCTCCAGCCTGTACCCCATCACAGTTTGGCCCG 1320
Db 387 -----SerGlyProProSerAlaMetThrPro-----Pro 396
QY 1321 TTCCCGCATCATCTCTTATCCAGAGCAGAGATTGTGAATCTTCTTCATCCCAACCCAG 1380
Db 397 TyrProGlnPheGluGln---SerGluThrGluThrValHisLeuPheIleProAlaLeu 415
QY 1381 GCTGTGGCGGCATCATCGGAAGAGGCGCACACATCAACAGCTGGCGAGATTTCGCC 1440
Db 416 SerValGlyAlaIleleGlyLysGlnGlyGlnHisIleLysGlnLeuSerArgPheAla 435
QY 1441 GGAGCCTCTATCAAGATTGCCCTTCGGAAGGCCAGACGTCAGCGAAAGATGCTCATC 1500
Db 436 GlyAlaSerIleLysIleAlaProAlaGluAlaProAspAlaLysValArgMetValIle 455
QY 1501 ATCACCGGCGCCAGCGAGCCAGTTCAGGCCCGGAGCGGATCTTTGGGAAGACTGAAA 1560
Db 456 IleThrGlyProProGluAlaGlnPheLysAlaGlnGlyArgIleTyrlGlyLysIleLys 475
QY 1561 GAGGAAACCTCTTTTAAACCCCAAGAAAGTGAAGCTGGAAGCGCATATCAGAGTGC 1620
Db 476 GluGluAsnPheValSerProLysGluGluValLysLeuGluAlaHisIleArgValPro 495
QY 1621 TCTTCCACAGTGGCCGGGTGATTGGCAAGAGGTGGCAAGACCGGTGAACGAACTCGAAC 1680
Db 496 SerPheAlaAlaGlyArgValIleGlyLysGlyLysThrValAsnGluLeuGlnAsn 515
QY 1681 TTAACAGTCGAGNAGTCATCGTCTCGTGACCAACAGCCAGATGAAATGAGGAAGTG 1740
```



Db 416 SerValGlyAlaIleIleGlyLysGlnGlyGlnHisIleLysGlnLeuSerArgPheAla 435  
QY 1441 GGAGCCTCTATCAAGNTTCCCTTGGGAGGCCAGAGCTCAGCGAAGAGTGTGTCATC 1500  
Db GlyAlaSerIleLysIleAlaProAlaGluAlaProAspAlaLysValArgMetValIle 455  
QY 1501 ATCACCAGGCGCCAGGAGCCAGTTCAGAGGCCAGGACGGATCTTTGGGAACTGAAA 1560  
Db 11eThrGlyProGluAlaGlnPheLysAlaGlnGlyArgIleGlyLysIleLys 475  
QY 1561 GAGGAAATCTTTAAACCCCAAGAGAGAGTGAAGCTGGAAGCCGATCAGAGTGC 1620  
Db 476 GluGluAsnPheValSerProLysGluGluValLysLeuGluAlaHisIleArgValPro 495  
QY 1621 TCTTCCACAGCTGGCGGTGATTGGCAAGGTGGCAGACCGTGAACGAACTGCAGAAC 1680  
Db SerPheAlaAlaGlyArgValIleGlyLysGlyLysThrValAsnGluLeuGlnAsn 515  
QY 1681 TPAACCACTGCAGAACTCATCTGCTCTGTGACCAACGCCAGATCAAAATCAGGAAGTG 1740  
Db LeuSerSerAlaGluValValProArgAspGlnThrProAspGluAsnGlnVal 535  
QY 1741 ATCTGCAAGTATATCGGCACCTCTTGTAGCCAGACTGCACAGCGCAAGATCAGGAAA 1800  
Db ValValLysIleThrGlyHisPheTyrAlaCysGlnValAlaGlnArgLysIleGlnGlu 555  
QY 1801 ATTGTACACAGTCAAGCAGCAGCAGCAGAAA 1833  
Db 11eLeuThrGlnValLysGlnHisGlnGlnGln 566

## RESULT 15

US-10-007-700-446  
; Sequence 446, Application US/10007700  
; Publication No. US20030064947A1

## GENERAL INFORMATION:

; APPLICANT: Wang, Tongtong  
; APPLICANT: Wang, Aijun  
; APPLICANT: Skeiky, Yasir A.W.  
; APPLICANT: Li, Samuel X.  
; APPLICANT: Kalos, Michael D.  
; APPLICANT: Henderson, Robert A.  
; APPLICANT: McNeill, Patricia D.  
; APPLICANT: Fanger, Neil  
; APPLICANT: Retter, Marc W.  
; APPLICANT: Durham, Margarita  
; APPLICANT: Fanger, Gary R.  
; APPLICANT: Vedwick, Thomas S.  
; APPLICANT: Carter, Darrick  
; APPLICANT: Watanabe, Yoshihiro  
; APPLICANT: Peckman, David W.  
; APPLICANT: Cai, Feng  
; APPLICANT: Foy, Teresa M.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER  
; FILE REFERENCE: 210121.455C17  
; CURRENT APPLICATION NUMBER: US/10/007,700  
; CURRENT FILING DATE: 2001-11-30  
; NUMBER OF SEQ ID NOS: 469  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 446  
; LENGTH: 579  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-007-700-446

## Alignment Scores:

Prod. No.: 2.67e-145 Length: 579  
Score: 1957.50 Matches: 387  
Percent Similarity: 79.19% Conservative: 81  
Best Local Similarity: 65.48% Mismatches: 94  
Query Match: 31.98% Indels: 29  
DB: 12 Gaps: 8

US-09-270-437D-6 (1-3412) x US-10-007-700-446 (1-579)  
QY 73 ATGAACAAGCTTTACATCGGAACCTTAGCCCGCGCTCACCCGCCAGACACTCCGGCAG 132  
Db 1 MetAsnLysLeuTyrIleGlyAsnLeuSerGluAsnAlaAProSerAspLeuGluSer 20  
QY 133 CTCCTTTGGGACAGAGCTGCCCTGGCGGACAGGTCTCTGCTGAAGTCCGCTACGCC 192  
Db 21 IlePheLysAspAlaLysIleProValSerGlyProPheLeuValLysThrGlyTyrAla 40  
QY 193 TTCGTGGGACTACCCGACGACAGAACTGGCCATCCCGCCATCGAGACCTCTCGGTTAA 252  
Db 41 PheValAspCysProAspGluSerIlePheLeuLysAlaIleGluAlaLeuSerGlyLys 60  
QY 253 CTGGAATTCATCGGAAAATCATGGAAGTTGATTACTAGTCTCTAAAAAGCTAAGGAGC 312  
Db 61 IleGluLeuHisGlyLysProIleGluValGluHisSerValProLysArgGlnArgIle 80  
QY 313 AGGAAATTCAGATTTCGAAACATCCTCTCTACCTGCGAGTGGAGGTGTGATGACTT 372  
Db 81 ArgLysLeuGlnIleArgAsnIleProHisLeuGlnTrpGluValLeuAspSerLeu 100  
QY 373 TTGGCTCAATATGGCAGACAGTGGAGAACTGTGAACAAGTCAACACAGACACAGAAACGCC 432  
Db 101 LeuValGlnTyrGlyValValGluSerCysGluGlnValAsnThrAspSerGlnThrAla 120  
QY 433 GTTGTCAACGTCAATATGCAACAGAGAAAGCAAAATAGCCATCGAAGAGCTAAGC 492  
Db 121 ValValAsnValThrTyrSerSerLysAspGlnAlaArgGlnAlaLeuAspLysLeuAsn 140  
QY 493 GGGCATCAGTTTGAGAACTACTCTTCAAGATTTCCTACATCCCGGATGAGAGGTGAGC 552  
Db 141 GlyPheGlnLeuGluAsnPheThrLeuLysValAlaTyrIleProAspGluThrAla 160  
QY 553 TCCCTTCGCCCCCTCAG-----CGAGCCAGAGCT-----GGGAGCACCTCTCCCGG 600  
Db 161 GlnGlnAsnProLeuGlnGlnProArgGlyArgArgGlyLeuGlyGlnArgGlySerSer 180  
QY 601 GAGCAAGGCCACGCCCTTGGGGGCATTTCTCAGGCCAGACAGATTGATTTCCTCCGCGG 660  
Db 181 ArgGlnGly---SerProGlySerValSerLysGlnLysProCysAspLeuProLeuArg 199  
QY 661 ATCTCTGCTCCCGCCAGCTTTGTTGGTCCCATCATCGGAAGAGGAGGCTTGAACATAAG 720  
Db 200 LeuLeuValProThrGlnPheValGlyAlaIleIleGlyLysGluGlyAlaThrIleArg 219  
QY 721 AACATCACTAAGCAGACCCAGTCCCGGTAGATATCCATAGAAAAGAGAACTCGGAGCT 780  
Db 220 AsnIleThrLysGlnThrGlnSerLysIleAspValHisArgLysGluAsnAlaGlyAla 239  
QY 781 GCAGAGAAGCTGTACCATCCATGCCACCCAGAGGGGACTTCTGAGCATGCGCGATG 840  
Db 240 AlaGluLysSerIleThrIleLeuSerThrProGluGlyThrSerAlaAlaCysLysSer 259  
QY 841 ATCTCTGAATCATCGAGAAGAGCGAGATGAGACCAACTAGCCGAGAGATTCTCTCTG 900  
Db 260 IleLeuGluIleMetHisLysGluAlaGlnAspIleLysPheThrGluGluIleProLeu 279  
QY 901 AAAATCTTGGCACACAATGGCTTGGTTGGAGAGCTGATTGAAAAGAGGAGGAGAAATTG 960  
Db 280 LysIleLeuAlaHisAsnAsnPheValGlyArgLeuIleGlyLysGluGlyArgAsnLeu 299  
QY 961 AAGAAATTTGAACATGAACAGGACCAAGATCAATCTCATCTTTGCGAGGATTTGAGC 1020  
Db 300 LysLysIleGluGlnAspThrAspThrLysIleThrIleSerProLeuGlnLeuThr 319  
QY 1021 ATATACACCGGAAACCAACATCATCTGTGAAGGCGACAGTTGAGGCTGTGCCAGTGT 1080  
Db 320 LeuTyrAsnProGluArgThrIleThrValLysGlyAsnValGluThrCysAlaLysAla 339  
QY 1081 GAGATAGAGATTATGAAGAAGCTGCGTGAGCGCTTTCAAAATATGATATGCTGGCTGTTAAC 1140  
Db 340 GluGluGluIleMetLysLysIleArgGluSerTyrGluAsnAspIleAlaSerMetAsn 359

```
QY 1141 CAACAGCCAACTCATCCAGGTTGAACCTCAGCGCATTTGGCATCTTTTCAACAGGA 1200
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
360 LeuGlnAlaHisLeuIleProGlyLeuAsnLeuAsnAlaLeuGlyLeuPhe----- 376
QY 1201 CTGTCCGTCTATCTCCACGACGAGGCCCCGCGAGCTCCCGCTGCCCTACCCAC 1260
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
377 -----ProThrSer-----GlyMetProProThr----- 386
QY 1261 CCCTTCACTACCCATCCGATCTCTCCAGCTGTACCCCATCACCAGTTTGGCCCG 1320
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
387 -----SerGlyProProSerAlaMetThrPro-----Pro 396
QY 1321 TTCCCGCATCATCTCTATCCAGCAGAGGATTGTGAATCTCTTCATCCCAACCCAG 1380
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
397 TyrProGlnPheGluGln---SerGluThrGluThrValHisLeuPheIleProAlaLeu 415
QY 1381 GCTGTGGGCGCCATCATCGGGAAGAGGGGCGACACATCAACAGCTGGCGAGATTCGCC 1440
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
416 SerValGlyAlaIleIleGlyGlnGlyGlnHisIleIleGlnLeuSerArgPheAla 435
QY 1441 GGAGCTCTCATCAAGATTCCCTCGGGAAGGCCAGACGTCAGCGAAAGGATGGTCATC 1500
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
436 GlyAlaSerIleIleIleAlaProAlaGluAlaProAspAlaLysValArgMetValIle 455
QY 1501 ATCAGCGGCGCCAGGAGCCAGTTCAAGGCCAGGCGGATCTTTGGGAAACTGAAA 1560
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
456 IleThrGlyProProGlnAlaGlnPheLysAlaGlnGlyArgIleIleGlyIleLys 475
QY 1561 GAGGAAACTCTTTAAACCCCAAGAGAGTGAAGCTGGAAGCGCATATCAGAGTSCCC 1620
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
476 GluGluAsnPheValSerProLysGluGluValLysLeuGluAlaHisIleArgValPro 495
QY 1621 TCTTCCACAGCTGCGCGGTGATTGCAAGAGGTGGCAAGCCGTGAACGAACTGCAGAAC 1680
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
496 SerPheAlaAlaGlyArgValIleGlyLysGlyLysThrValAsnGluLeuGlnAsn 515
QY 1681 TTAACCAAGTCAGAGTCATCGTCCTCGTGACCAACGCCAGATCAAAATGAGGAAGTG 1740
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
516 LeuSerSerAlaGluValValProArgAspGlnThrProAspGluAsnAspGlnVal 535
QY 1741 ATCGTCAGAAATTATCGGCACTTTCTTTGCTAGCCAGACTGCACAGCGCAAGATCAGGGAA 1800
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
536 ValValLysIleThrGlyHisPheTyrAlaCysGlnValAlaGlnArgLysIleGlnGlu 555
QY 1801 ATTGTACAACAGTGAAGCAGCAGGACGAGAAA 1833
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
556 IleLeuThrGlnValLysGlnHisGlnGlnGln 566
```

Search completed: July 16, 2004, 10:57:42

Job time : 357 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: July 16, 2004, 10:47:12 ; Search time 149 Seconds  
(without alignments)

12940.296 Million cell updates/sec

Title: US-09-270-437D-6

Perfect score: 1120

Sequence: 1 ggcagcgaggagcgagga.....aaccttgaaaatgtttattt 3412

Scoring table:

OLIGO  
Xgapop 60.0 , Xgapext 60.0  
Ygapop 60.0 , Ygapext 60.0  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 1586107 seqs, 282547505 residues

Word size: 1

Total number of hits satisfying chosen parameters: 2987630

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

--MODEL=frame+n2p.model -DEV=xlp  
-Q=/cgn2\_1/USPRO\_spool\_p/US09270437/runat 16072004 113126 13935/app query.fasta 1.3591  
-DB=A Geneseq 29Jan04 -QFMT=fastan -SUFFIX=Oligo.rag -MINMATCH=0.1 -LoopCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS-human40.cdi  
-LIST=45 -DOCALIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEA=SIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USRF=US09270437.cgn\_1\_1\_253\_@runat 16072004 113126 13935 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOPOP=6  
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : A Geneseq 29Jan04.\*  
1: Geneseqp1980s.\*  
2: Geneseqp1990s.\*  
3: Geneseqp2000s.\*  
4: Geneseqp2001s.\*  
5: Geneseqp2002s.\*  
6: Geneseqp2003as.\*  
7: Geneseqp2003bs.\*  
8: Geneseqp2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	603	53.8	614	4	ABG06794 Novel hum
2	511	45.6	620	4	AAU16163 Human nov
3	511	45.6	620	6	ABU5232 Human nov
4	378	33.8	594	4	ABG06795 Novel hum
5	357	31.9	556	5	ABG96346 Human ova
6	160	14.3	171	4	AAU16166 Human nov
7	160	14.3	171	6	ABU5235 Human nov
8	97	8.7	97	3	AAG03261 Human sec
9	75	6.7	192	4	AAU16164 Human nov
10	75	6.7	192	6	ABU5233 Human nov

11	66	5.9	171	4	AAU16583 Human nov
12	66	5.9	171	6	ABU55652 Human nov
13	44	3.9	555	6	ABU89799 Novel hum
14	32	2.9	187	6	AAO23971 Human IGF
15	20	1.8	47	4	AAU16581 Human nov
16	20	1.8	47	6	ABU55650 Human nov
17	19	1.7	148	4	ABG21962 Novel hum
18	19	1.7	250	4	AAU16579 Human nov
19	19	1.7	250	6	ABU55648 Human nov
20	19	1.7	261	4	AAU16161 Human nov
21	19	1.7	261	6	ABU55230 Human nov
22	19	1.7	339	4	AAM53826 Human pol
23	19	1.7	577	2	AAY30649 A murine
24	19	1.7	579	3	AAU11328 Human lun
25	19	1.7	579	3	AAB11365 Human lun
26	19	1.7	579	5	ABU74960 Human lun
27	19	1.7	579	5	ABU75053 Human lun
28	19	1.7	579	5	ABU74997 Human lun
29	19	1.7	579	5	ABU75054 Human lun
30	19	1.7	579	5	ABU61917 Human lun
31	19	1.7	579	5	ABU61974 Human lun
32	19	1.7	579	5	ABU61880 Human lun
33	19	1.7	579	5	ABU61973 Human lun
34	19	1.7	579	7	ADA28536 Recombina
35	19	1.7	579	7	ADA28539 Recombina
36	19	1.7	579	7	ADA28438 Human lun
37	19	1.7	579	7	ADA28266 Human lun
38	19	1.7	579	7	ADD14066 Human src
39	19	1.7	579	7	ADE53471 Human lun
40	19	1.7	586	5	ABU75048 Human lun
41	19	1.7	586	5	ABU75048 Human lun
42	19	1.7	586	7	ADA28517 Recombina
43	18	1.6	619	4	ABG21963 Novel hum
44	17	1.5	583	4	ABG12592 Novel hum
45	15	1.3	20	5	ABU75062 Human lun

#### ALIGNMENTS

RESULT 1  
ABG06794

ID ABG06794 standard; protein; 614 AA.

XX ABG06794;

XX DT 13-FEB-2002 (first entry)

XX DE Novel human diagnostic protein #6785.

XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;

XX KW food supplement; medical imaging; diagnostic; genetic disorder.

XX OS Homo sapiens.

XX PN WO200175067-A2.

XX PD 11-OCT-2001.

XX PF 30-MAR-2001; 2001WO-US008631.

XX PR 31-MAR-2000; 2000US-00540217.

XX PR 23-AUG-2000; 2000US-00649167.

XX (HYSE-) HYSEQ INC.

XX Drmanac RT, Liu C, Tang YT;

XX WPI; 2001-639362/73.

XX N-PSDB; AAS70981.

XX New isolated polynucleotide and encoded polypeptides, useful in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits and to assess

PT biodiversity.  
 XX Claim 20; SEQ ID NO 37153; 103pp; English.  
 XX The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. ABG0010-ABG0377 represent novel human diagnostic amino acid sequences of the invention. Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

XX SQ Sequence 614 AA;  
 Alignment Scores:  
 Pred. No.: 0 Length: 614  
 Score: 603.00 Matches: 603  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 53.84% Indels: 0  
 DB: 4 Gaps: 0

US-09-270-437d-6 (1-3412) x ABG06794 (1-614)

QY 58 GGGAGAGACGGATGATGAACAGCTTTACATCGGGAACCTGAGCCCGCGCTCACGGC 117  
 Db 12 GlyLysArgMetMetAsnLysLeuTyrIleGlyAsnLysSerProAlaValThrAla 31  
 QY 118 GACGACCTCCGACGCTCTTTGGGACAGGAGCTGCCCTCGCGGACAGCTCTCTGCTG 177  
 Db 32 AspAspLeuArgGlnLeuPheGlyAspArgLysLeuProLeuAlaGlyGlnValLeuLeu 51  
 QY 178 AAGTCGGCTACGCTCTCTGACTACCCCGACCAAGACTGGGCCATCCGCGCCATCGAG 237  
 Db 52 LysSerGlyTyrAlaPheValAspTyrProAspGlnAsnTyrAlaIleArgAlaIleGlu 71  
 QY 238 ACCCTTCGGTAAAGTGAATTGCATGGGAAATCATGGAAGTTGATTACTCAGTCTCT 297  
 Db 72 ThrLeuSerGlyLysValGluLeuHisGlyLysIleMetGluValAspTyrSerValSer 91  
 QY 298 AAAAAGCTAAGACGAGGAAATTCAGATTGCAACATCCCTCCTCACCTGAGTGGGAG 357  
 Db 92 LysLysLeuArgSerArgLysIleGlnIleArgAsnIleProProHisLeuGlnTyrGlu 111  
 QY 358 GTGTTGGATGGACTTTTGGCTCAATATGGGACAGTGGAGAATGTGGAACAAGTCAACA 417  
 Db 112 ValLeuAspGlyLeuLeuAlaGlnTyrGlyThrValGluAsnValGluGlnValAsnThr 131  
 QY 418 GACACGAACCGCGTCTGCAAGTCACATATGCAACAGGAGGAGAAATAGCC 477  
 Db 132 AspThrGluThrAlaValValAsnValThrTyrAlaThrArgGluGluAlaLysIleAla 151  
 QY 478 ATGGAGAAGCTAAGCGGGCATCAGTTTGAGAACTACTCTCTCAAGATTTCTACATCCCG 537  
 Db 152 MetGluLysLeuSerGlyHisGlnPheGluAsnTyrSerPheLysIleSerTyrIlePro 171  
 QY 538 GATGAAGAGGTGAGCTCCCTTCGCCCCCTCAGCGAGCCGAGCTGGGACCACTTCC 597  
 Db 172 AspGluGluValSerSerProSerProGlnArgAlaGlnArgGlyAspHisSerSer 191

QY 598 CGGAGCAAGCCACGCCCCCTGGGGGCACCTTCTCAGGCCACAGATTGATTTCCGCTG 657  
 Db 192 ArgGluGlnGlyHisAlaProGlyGlyThrSerGlnAlaArgGlnIleAspPheProLeu 211  
 QY 658 CGGATCTGGTCCCGCCACCCAGTTTGTGTCCTCATCGAAAGGAGGCTTGACCATTA 717  
 Db 212 ArgIleLeuValProThrGlnPheValGlyAlaIleIleGlyLysGluGlyLeuThrIle 231  
 QY 718 AAGAACATCACTAAGCAGACCCAGTCCCGGTAGATATCCATAGAAAAGAGAACTCTCGA 777  
 Db 232 LysAsnIleThrLysGlnThrGlnSerArgValAspIleHisArgLysGluAsnSerGly 251  
 QY 778 GCTGAGAGAAGCCCTGTCCATCCATCCATCCACCCAGAGGGAGCTTCTGAGACATCCGC 837  
 Db 252 AlaAlaGluLysProValThrIleHisAlaThrProGluGlyThrSerGluAlaCysArg 271  
 QY 838 ATGATCTTTGAAATCATGCAAGAGAGGACAGATGAGCAAACTAGCCGAGAGATTCT 897  
 Db 272 MetIleLeuGluIleMetGlnLysGluAlaAspGluThrLysLeuAlaGluGluLeuPro 291  
 QY 898 CTGAAAATCTTTGGCACACAATGGCTTGGTGGAGACTGATTTGAAAAGAGAGCGAGAAT 957  
 Db 292 LeuLysIleLeuAlaHisAsnGlyLeuValGlyArgLeuIleGlyLysGluGlyArgAsn 311  
 QY 958 TTGAGAAAATTTGAACATGAACAGGACAGGACCAAGATAACAATCTCATCTTTGAGATTG 1017  
 Db 312 LeuLysLysIleGluHisGluThrGlyThrLysIleThrIleSerSerLeuGlnAspLeu 331  
 QY 1018 AGCATATACAAACCCGGAAGAACCATCACTGTGAAGGCGACAGTTGAGCCCTGTGCCAGT 1077  
 Db 332 SerIleTyrAsnProGluArgThrIleThrValLysGlyThrValGluAlaCysAlaSer 351  
 QY 1078 GCTGAGATAGAGATTATGAAGAGCTGCGTGAAGGCTTTGAAAATGATATGCTGCTGTT 1137  
 Db 352 AlaGluIleGluIleMetLysLysLeuArgGluAlaPheGluAsnAspMetLeuAlaVal 371  
 QY 1138 AACCAACAGCAATCTGATCCCGGTTGAACCTCAGCGACCTGAGCTCTTTTCAACA 1197  
 Db 372 AsnGlnGlnAlaAsnLeuIleProGlyLeuAsnLeuSerAlaLeuGlyIlePheSerThr 391  
 QY 1198 GGACTGTGCTGCTATCTCCACAGAGGCGCCCGGAGCTCCCGCTCCCGCTAC 1257  
 Db 392 GlyLeuSerValLeuSerProProAlaGlyProArgGlyAlaProProAlaAlaProTyr 411  
 QY 1258 CACCCCTTCACTACCCACTCCGGTACTTCTCCAGCTGTACCCCATCACCAGTTGCG 1317  
 Db 412 HisProPheThrThrHisSerGlyTyrPheSerSerLeuTyrProHisGlnPheGly 431  
 QY 1318 CGTTTCCGCGCATCATCTCTTATCCAGACGAGAGATTGAAATCTCTTCATCCCAAGC 1377  
 Db 432 ProPheProHisHisSerTyrProGluGlnGluIleValAsnLeuPheIleProThr 451  
 QY 1378 CAGGCTGTGGGCGCATCATCGGGAAGAGGGGGCACACATCAACAGCTGGCGAGATTTC 1437  
 Db 452 GlnAlaValGlyAlaIleIleGlyLysLysGlyAlaHisIleLysGlnLeuAlaArgPhe 471  
 QY 1438 GCGGAGCTCTTCAAGATTCCCTCGGAGGCCAGAGCTGAGCTGAGAGGATGCTC 1497  
 Db 472 AlaGlyAlaSerIleLysIleAlaProAlaGluGlyProAspValSerGluArgMetVal 491  
 QY 1498 ATCATCACCGGCGCCAGGAGCCAGTTTCAAGGCCGAGGACGGATCTTTTGGGAAACTG 1557  
 Db 492 IleIleThrGlyProProGluAlaGlnPheLysAlaGlnGlyArgIlePheGlyLysLeu 511  
 QY 1558 AAAGAGGAAAATCTTTTAAACCCCAAGAGAGTGAAGCTGAGCGCATATCAGATG 1617  
 Db 512 LysGluGluAsnPhePheAsnProLysGluGluValLysLeuGluAlaHisIleArgVal 531  
 QY 1618 CCTCTTCCACAGCTGGCGGCTGATTGGCAAGGTGGCAAGACCGTGAACCACTGCAG 1677  
 Db 532 ProSerSerThrAlaGlyArgValIleGlyLysGlyLysThrValAsnGluLeuGln 551



PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249219P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249264P.  
PR 17-NOV-2000; 2000US-0249265P.  
PR 17-NOV-2000; 2000US-0249297P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 17-NOV-2000; 2000US-0249300P.  
PR 01-DEC-2000; 2000US-0250160P.  
PR 01-DEC-2000; 2000US-0250391P.  
PR 05-DEC-2000; 2000US-0251030P.  
PR 05-DEC-2000; 2000US-0251988P.  
PR 05-DEC-2000; 2000US-0256719P.  
PR 06-DEC-2000; 2000US-0251479P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.  
PR 08-DEC-2000; 2000US-0251989P.  
PR 08-DEC-2000; 2000US-0251990P.  
PR 11-DEC-2000; 2000US-0254097P.  
PR 05-JAN-2001; 2001US-0259678P.  
XX  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX

PI Rosen CA, Barash SC, Ruben SM;

XX WPI; 2001-488783/53.

XX N-PSDB; AAS26150.

XX New nucleic acid molecules encoding 461 human secreted proteins for  
PT diagnosing, preventing, treating or ameliorating medical conditions and  
PT used as food additives or preservatives.

XX Claim 11; SEQ ID NO 1116; 980pp; English.

XX The invention relates to isolated nucleic acid molecules and their  
CC encoded secreted proteins. The nucleic acids and proteins are used to  
CC prevent, treat or ameliorate a medical condition in e.g. humans, mice,  
CC rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used  
CC in diagnosing a pathological condition or susceptibility to a  
CC pathological condition. Antibodies to the proteins can also be used in  
CC alleviating symptoms associated with the disorders and in diagnostic  
CC immunoassays e.g. radioimmunoassays or enzyme linked immunosorbent assays  
CC (ELISA). Disorders which are diagnosed or treated include autoimmune  
CC diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g.  
CC neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac  
CC arrest, cerebrovascular disorders e.g. cerebral ischaemia, angiogenesis,  
CC nervous system disorders e.g. Alzheimer's disease, infections caused by  
CC bacteria, viruses and fungi and ocular disorders e.g. corneal infection,  
CC and many other disorders listed in the specification. The polypeptides  
CC can also be used to aid wound healing and epithelial cell proliferation,  
CC to prevent skin aging due to sunburn, to maintain organs before  
CC transplantation, for supporting cell culture of primary tissues, to  
CC regenerate tissues and in chemotaxis. The polypeptides can also be used  
CC as a food additive or preservative to increase or decrease storage  
CC capabilities, fat content, lipid, protein, carbohydrate, vitamins,  
CC minerals, cofactors and other nutritional components. The present  
CC sequence represents a novel secreted protein of the invention. Note: The  
CC sequence data for this patent did not form part of the printed

Alignment Scores:

Pred. No.: 0 Length: 620  
Score: 511.00 Matches: 511  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 45.62% Indels: 0  
DB: 4 Gaps: 0

US-09-270-437D-6 (1-3412) x AAUI6163 (1-620)

QY 70 ATGATGAACAAAGCTTTATCATCGGAAACCTGAGCCGCCGCTCACGCCGACGACCTCCGG 129  
Db 22 MetMetAsnLysLeuTyrIleGlyAsnLeuSerProAlaValThrAlaAspLeuArg 41  
QY 130 CAGCTCTTTGGGACAGAGAGCTGCCCTGGGGACAGGTCTCTGCTCAAGTCGGCTAC 189  
Db 42 GlnLeuPheGlyAspArgLysLeuProLeuAlaGlyGlnValLeuLeuLysSerGlyTyr 61  
QY 190 GCCTTCGTGGACTACCCCGACAGAACTGGGCGCATCCGCGCATCGAGACCTCTCGGT 249  
Db 62 AlaPheValAspTyrProAspGlnAsnTrpAlaIleArgAlaIleGluThrLeuSerGly 81  
QY 250 AAAGTGAATTTCATGCGAAATCATGGAAGTTGATTACTCAGTCTCTAAAAAGTAAG 309  
Db 82 LysValGluLeuHisGlyLysIleMetGluValAspTyrSerValSerLysLysLeuArg 101  
QY 310 AGCAGGAAATTCAGATTCGAAACATCCCTCTCACCTGCAGTGGAGGTGTTCATGGA 369  
Db 102 SerArgLysIleGlnIleArgAsnIleProHisLeuGlnTrpGluValLeuAspGly 121  
QY 370 CTTTGGCTCAATATGGACAGTGGGAATGTGGAACAAGTCAACACACACAGAACCC 429  
Db 122 LeuLeuAlaGlnTyrGlyThrValGluAsnValGluGlnValAsnThrAspThrGluThr 141  
QY 430 GCCGTTGTCACGTCACATATGCAACAGAGAGAGCAAAATAGCCATGAGAGCTA 489  
Db 142 AlaValValAsnValThrTyrAlaThrArgGluGluAlaLysIleAlaMetGluLysLeu 161  
QY 490 AGCGGSCATCAGTTTGAAGACTACTCTTCAAGATTTCTACATCCCGATGAAGAGGTG 549  
Db 162 SerGlyHisGlnPheGluAsnTyrSerPheLysIleSerTyrIleProAspGluGluVal 181  
QY 550 AGCTCCCTTCGCCCTTCAGCAGCCCGAGCTGGGACCACTCTTCCCGGAGCAGAGC 609  
Db 182 SerSerProSerProGlnArgAlaGlnArgGlyAspHisSerSerArgGluGlnGly 201  
QY 610 CAGGCCCTTCGGGGACATCTCTCAGGCCACAGATTGATTTCCCGTCGGGATCCTGTC 669  
Db 202 HisAlaProGlyGlyThrSerGlnAlaArgGlnIleAspPheProLeuArgIleLeuVal 221  
QY 670 CCACCCAGTTTGTGGTCCCATCATCGGAAAGGAGGGCTTGACCATAAAGAACATCACT 729  
Db 222 ProThrGlnPheValGlyAlaIleIleGlyLysGluGlyLeuThrIleLysAsnIleThr 241  
QY 730 AAGCAGACCCAGTCCGGGTAGATATCCATAGAAAGAGAACTCTGGAGCTGCAGACAG 789  
Db 242 LysGlnThrGlnSerArgValAspIleHisArgLysGluAsnSerGlyAlaAlaGluLys 261  
QY 790 CCTGTCCACATCCATGCCCCCGAGGGAGCTTCTGAAGCATGCCGCGATGATTTCTTGA 849  
Db 262 ProValThrIleHisAlaThrProGluGlyThrSerGluAlaCysArgMetIleLeuGlu 281  
QY 850 ATCATCCAGAAAGAGGCAGATGAGACCAACTAGCCGAAGAGATTCCTCTGAAATCTTG 909  
Db 282 IleMetGlnLysGluAlaAspGluThrLysLeuAlaGluGluIleProLeuLysIleLeu 301  
QY 910 GCACACAATGGCTTGGTGGAGACTGATTGAAAGAGAGGACAAATTTGAGAAATTT 969  
Db 302 AlaHisAsnGlyLeuValGlyArgLeuIleGlyLysGluGlyArgAsnLeuLysLysIle 321  
QY 970 GAACATGAACACAGGGACCAAGATAACAATCTCATCTTTTCAGGATTTGACATATCAAC 1029  
Db 322 GluHisGluThrGlyThrLysIleThrIleSerSerSerLeuGlnAspLeuSerIleTyrAsn 341



QY 1030 CCGGAAAGAACCATCTGCTGAAGCGGCACAGTGTGAGGCGCTGTGCAGTGTGCAGATAGAG 1089  
Db |||||  
342 ProGluArgThrIleThrValLysGlyThrValGluAlaCysAlaSerAlaGluIleGlu 361  
QY 1090 ATTATGAAGAAGCTCGTGAAGCGCTTTGAAATGATATGCTGGCTGTAAACCAACAGCC 1149  
Db |||||  
362 IleMetLysLysLeuArgGluAlaPheGluAsnAspMetLeuAlaValAsnGlnGlnAla 381  
QY 1150 AATCTGATCCAGGTTGAACCTCAGCGCATCTGGCATCTTTCAACAGGACTGTCGCTG 1209  
Db |||||  
382 AsnLeuIleProGlyLeuAsnLeuSerAlaLeuGlyIlePheSerThrGlyLeuSerVal 401  
QY 1210 CTATCTCCACAGAGCGGCCCGGAGCTCCCGCGCTCCACCATCACCAGTTTGGCCCGCTCCCGCAT 1259  
Db |||||  
402 LeuSerProProAlaGlyProArgGlyAlaProProAlaAlaProTyrHisProPheThr 421  
QY 1270 ACCCACTCCGATACTTCTCCAGCGCTGTACCCCATCACCAGTTTGGCCCGCTCCCGCAT 1329  
Db |||||  
422 ThrHisSerGlyTyrPheSerSerLeuTyrProHisHisGlnPheGlyProPheProHis 441  
QY 1330 CATCACTTTATCCAGACGAGATGTGAATCTCTTCATCCCAACCCAGGCTGTGGGC 1389  
Db |||||  
442 HisHisSerTyrProGluGlnGluIleValAsnLeuPheIleProThrGlnAlaValGly 461  
QY 1390 GCCATCATCGGAAGAAGGGGCGACATCAACAGCTGGCAGATTCCCGGAGCCTCT 1449  
Db |||||  
462 AlaIleIleGlyLysLysGlyAlaHisIleLysGlnLeuAlaArgPheAlaGlyAlaSer 481  
QY 1450 ATCAAGATTGCCCTCGGAGGCGCCAGCTCAGCGAAAGGATGTCATCATCACCGGG 1509  
Db |||||  
482 IleLysIleAlaProAlaGluGlyProAspValSerGluArgMetValIleIleThrGly 501  
QY 1510 CCACCGAAGCCAGTTCAAGGCCCGAGCGGATCTTTGGGAAACTGAAAGAGGAAAC 1569  
Db |||||  
502 ProProGluAlaGlnPheLysAlaGlnGlyArgIlePheGlyLysLeuLysGluGluAsn 521  
QY 1570 TTCTTTAACCCCAAGAGAGTGAAGCTGGA 1602  
Db |||||  
522 PhePheAsnProLysGluGluValLysLeuGlu 532  
RESULT 3  
ID ABU55232 standard; protein; 620 AA.  
XX AC ABU55232;  
XX AC  
XX AC  
DT 18-MAR-2003 (first entry)  
XX XX  
DE Human novel polypeptide #319.  
XX KW Human; neural disorder; immune system disorder; renal disorder;  
KW muscular disorder; respiratory disease; reproductive disorder;  
KW gastrointestinal disorder; pulmonary disorder; cardiovascular disorder;  
KW hyperproliferative disorder; inflammatory disease; allergic reaction;  
KW blood related disorder; cancer; immunosuppressive; antiinflammatory;  
KW cardiovascular; nephrotropic; cytostatic; antiallergic; thrombolytic;  
XX KW haemostatic; antiarteriosclerotic.  
OS Homo sapiens.  
XX PN US2002132753-A1.  
XX PD 19-SEP-2002.  
XX XX  
XX PF 17-JAN-2001; 2001US-00764864.  
XX PR 31-JAN-2000; 2000US-0179065P.  
PR 04-FEB-2000; 2000US-0180628P.  
PR 28-JUN-2000; 2000US-0214986P.  
PR 07-JUL-2000; 2000US-0216647P.  
PR 07-JUL-2000; 2000US-0216880P.  
PR 11-JUL-2000; 2000US-0217487P.

PR 11-JUL-2000; 2000US-0217496P.  
PR 14-JUL-2000; 2000US-0218290P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 26-JUL-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225447P.  
PR 14-AUG-2000; 2000US-0225757P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 22-AUG-2000; 2000US-0226868P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239335P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.  
XX (ROSE/) ROSEN C A.  
PA (RUBE/) RUBEN S M.  
PA (BARA/) BARASH S C.  
XX Rosen CA, Ruben SM, Barash SC;  
PI WPI; 2003-147444/14.  
XX N-PSDB; ABX73491.  
XX New polypeptides and nucleic acids, useful in gene therapy for treating,  
PT inhibiting or preventing e.g. neural, immune system, muscular,  
PT respiratory, reproductive, gastrointestinal, pulmonary, cardiovascular or  
PT renal disorders.  
XX Claim 11; SEQ ID NO 1116; 402pp; English.  
PS The invention relates to human novel polypeptides and their associated  
XX polynucleotides. The polypeptides and polynucleotides are useful in gene  
CC therapy for treating, inhibiting or preventing neural disorders, immune  
CC system disorders (e.g. systemic lupus erythematosus, rheumatoid arthritis  
CC and multiple sclerosis), muscular disorders, respiratory diseases (e.g.  
CC nasal vestibulitis, nasal polyps and sinusitis), reproductive disorders,  
CC gastrointestinal disorders, pulmonary disorders, cardiovascular disorders  
CC (e.g. congenital heart defects, Ebstein's anomaly and hypoplastic left  
CC heart syndrome), renal disorders (e.g. acute kidney failure and end-stage  
CC renal disease), hyperproliferative disorders (e.g. Hodgkin's disease and  
CC leukaemia), inflammatory diseases (e.g. septic shock, bursitis and  
CC appendicitis), allergic reactions and conditions (e.g. asthma), blood

CC related disorders (e.g. thrombosis, atherosclerosis and myocardial  
 CC infarction) and cancerous diseases. Sequences ABU54914-ABU55699 and  
 CC ABU55748 represent human novel polypeptides of the invention  
 XX  
 SQ Sequence 620 AA;

Alignment Scores:  
 Pred. No.: 0 Length: 620  
 Score: 511.00 Matches: 511  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 45.62% Indels: 0  
 DB: 6 Gaps: 0

US-09-270-437D-6 (1-3412) x ABU55232 (1-620)

QY 70 ATGATGAACAAGCTTTACATCGGGAACCTGAGCCCGCGCTCACCGCGACGACCTCCGG 129  
 Db 22 MetMetAsnLysLeuTyriLeGlyAsnLeuSerProAlaValThrAlaAspAspLeuArg 41  
 QY 130 CAGCTCTTTGGGACAGGAAGCTGCCCTGGGGGACAGGTCTCTGCTGAAGTCCGGCTAC 189  
 Db 42 GlnLeuPheGlyAspArgLysLeuProLeuAlaGlyGlnValLeuLeuLysSerGlyTyr 61  
 QY 190 GCCTTCGTGGACTACCCCGACAGAACTGGGCGCATCCGCGCCATCGAGACCTCTCGGT 249  
 Db 62 AlaPheValAspTyrProAspGlnAsnTrpAlaIleArgAlaIleGlnThrLeuSerGly 81  
 QY 250 AAAGTGGGAATTCATGGGAAATCATGGAAGTTGATTCTCAGTCTCTAAAAGCTAAAG 309  
 Db 82 LysValGluLeuHisGlyLysIleMetGluValAspTyrSerValSerLysLysLeuArg 101  
 QY 310 AGCAGAAATTCAGATTCGAACATCCCTCTCCTACCTCAGTGGAGGTGTGATGGA 369  
 Db 102 SerArgLysIleGlnIleArgAsnIleProProHisLeuGlnTrpGluValLeuAspGly 121  
 QY 370 CTTTGGCTCAATATGGACAGTGGAGATGGAACAGTCAACACAGACACAGAAAC 429  
 Db 122 LeuLeuAlaGlnTyrGlyThrValGluAsnValGluGlnValAsnThrAspThrGluThr 141  
 QY 430 GCCGTGTCAACGTCACATATCACACAGAGAGAGCAAAATAGCCATGGAGAGCTA 489  
 Db 142 AlaValValAsnValThrTyrAlaThrArgGluGluAlaLysIleAlaMetGluLysLeu 161  
 QY 490 ACGGGCATCAGTTTGGACACTACTCTTCAAGATTCTCTACATCCCGATGAGAGGTG 549  
 Db 162 SerGlyHisGlnPheGluAsnTyrSerPheLysIleSerTyrIleProAspGluGluVal 181  
 QY 550 AGCTCCCTTCGCCCTCAGCGAGCCCGAGCTGGGACCACTCTTCCCGGAGCAAGGC 609  
 Db 182 SerSerProSerProGlnArgAlaGlnArgGlyAspHisSerSerArgGluGlnGly 201  
 QY 610 CACGCCCTTGGGGCACTTCTCAGGCCACAGATTGATTTCCGCTGGGATCTCTGGTC 669  
 Db 202 HisAlaProGlyGlyThrSerGlnAlaArgGlnIleAspPheProLeuArgIleLeuVal 221  
 QY 670 CCACCCAGTTTCTGGTCCCATTCGGAAGAGGGCTTGACCAATAAGACATCACT 729  
 Db 222 ProThrGlnPheValGlyAlaIleIleGlyLysGluGlyLeuThrIleLysAsnIleThr 241  
 QY 730 AAGCAGACCCAGTCCCGGGTAGATATCCATAGAAAAGAGAACTCTGGAGCTGCAGAGAAG 789  
 Db 242 LysGlnThrGlnSerArgValAspIleHisArgLysGluAsnSerGlyAlaAlaGluLys 261  
 QY 790 CTGTGCATTCATCCACCCAGAGGGGACTTCTGAGCATGCCGATGATCTTGAA 849  
 Db 262 ProValThrIleHisAlaThrProGluGlyThrSerGluAlaCysArgMetIleLeuGlu 281  
 QY 850 ATCATCAGAAAGAGGACAGATGAGCCAAACTAGCCGAAGAGATTCTCTGAAAATCTTG 909  
 Db 282 IleMetGlnLysGluAlaAspGluThrLysLeuAlaGluGluIleProLeuLysIleLeu 301  
 QY 910 GCACACAATGGCTTGGTGAAGACTGATTGGAAAAGAGGACAGAAATTTGAAGAAAT 969

Db 302 AlaHisandGlyLeuValGlyArgLeuIleGlyLysGluGlyArgAsnLeuLysLysIle 321  
 QY 970 GAACATGAAACAGGACCAAGATAACAATCTCATCTTTCAGGATTTGAGCATATACAC 1029  
 Db 322 GluHisGluThrGlyThrLysIleThrIleSerSerLeuGlnAspLeuSerIleTyrAsn 341  
 QY 1030 CCGGAAAGAACCATCACTGTGAAGGCGACAGCTTCAGGCCTGTCCAGTGTGAGATAGAG 1089  
 Db 342 ProGluArgThrIleThrValLysGlyThrValGluAlaCysAlaSerAlaGluIleGlu 361  
 QY 1090 ATTATGAAGAGCTGCGTGAAGCCCTTTGAAAATGATGTGCTGTGTAAACCAAGACC 1149  
 Db 362 IleMetLysLysLeuArgGluAlaPheGluAsnAspMetLeuAlaValAsnGlnGlnAla 381  
 QY 1150 AATCTGATCCAGGTTGAACCTCAGCGCCTTCAGGCATCTTTCAACAGGACTGTCGGTG 1209  
 Db 382 AsnLeuIleProGlyLeuAsnLeuSerAlaLeuGlyIlePheSerThrGlyLeuSerVal 401  
 QY 1210 CTATCTCCACAGCAGGCGCCCGGAGCTCCCGCGCTGCCCTACACCCCTTCACT 1269  
 Db 402 LeuSerProProAlaGlyProArgGlyAlaProProAlaAlaProTyrHisProPheThr 421  
 QY 1270 ACCCATCTCGGATCTTCTCCAGCCTGTACCCCATCACAGTTTGGCCCGCTTCCGCAAT 1329  
 Db 422 ThrHisSerGlyTyrPheSerSerLeuTyrProHisGlnPheGlyProPheProHis 441  
 QY 1330 CATCACTTTATCCAGAGCAGGAGATTGTGATCTCTCATCCCAACCCAGGCTGTGGC 1389  
 Db 442 HisHisSerTyrProGluGlnGluIleValAsnLeuPheIleProThrGlnAlaValGly 461  
 QY 1390 GCATCATCGGGAAGAGGGGCGACACATCAAAAGCTGGCGAGATTCCCGGAGGCTCT 1449  
 Db 462 AlaIleIleGlyLysLysGlyAlaHisIleLysGlnLeuAlaArgPheAlaGlyAlaSer 481  
 QY 1450 ATCAAGATTGCCCTCGCGAAGCCCGACAGTCAAGGAAAGGATGGTCTATCATCCCGG 1509  
 Db 482 IleLysIleAlaProAlaGluGlyProAspValSerGluArgMetValIleIleThrGly 501  
 QY 1510 CCACCGGAGGCCAGTTCGAAGCCCGAGGCGAGTCTTGGCAACTGAAAGAGGAAAC 1569  
 Db 502 ProProGluAlaGlnPheLysAlaGlnGlyArgIlePheGlyLysLeuLysGluGluAsn 521  
 QY 1570 TTCTTTAAACCCCAAGAGAGAGTGAAGCTGGAA 1602  
 Db 522 PhePheAsnProLysGluGluValLysLeuGlu 532

RESULT 4  
 ABG06795  
 ID ABG06795 standard; protein; 594 AA.  
 XX  
 AC ABG06795;  
 XX  
 DT 13-FEB-2002 (first entry)  
 XX  
 DE Novel human diagnostic protein #6786.  
 XX  
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
 KW food supplement; medical imaging; diagnostic; genetic disorder.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200175067-A2.  
 XX  
 XX 11-OCT-2001.  
 PD  
 XX 30-MAR-2001; 2001WO-US008631.  
 PF  
 XX 31-MAR-2000; 2000US-00540217.  
 PR  
 PR 23-AUG-2000; 2000US-00649167.  
 XX  
 PA (HYSB-) HYSEQ INC.  
 XX

PI Drmanac RT, Liu C, Tang YT;

XX WPI; 2001-639362/73.  
DR N-PSDB; AAS70982.

XX New isolated polynucleotide and encoded polypeptides, useful in  
PT diagnostics, forensics, gene mapping, identification of mutations  
PT responsible for genetic disorders or other traits and to assess  
XX biodiversity.

PS Claim 20; SEQ ID NO 37154; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and polypeptide (II)  
CC sequences. (I) is useful as hybridisation probes, polymerase chain  
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
CC and in recombinant production of (II). The polynucleotides are also used  
CC in diagnostics as expressed sequence tags for identifying expressed  
CC genes. (I) is useful in gene therapy techniques to restore normal  
CC activity of (II) or to treat disease states involving (II). (II) is  
CC useful for generating antibodies against it, detecting or quantitating a  
CC polypeptide in tissue, as molecular weight markers and as a food  
CC supplement. (II) and its binding partners are useful in medical imaging  
CC of sites expressing (II). (I) and (II) are useful for treating disorders  
CC involving aberrant protein expression or biological activity. The  
CC polypeptide and polynucleotide sequences have applications in  
CC diagnostics, forensics, gene mapping, identification of mutations  
CC responsible for genetic disorders or other traits to assess biodiversity  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic  
CC patent did not appear in the invention. Note: The sequence data for this  
CC electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences

XX Sequence 594 AA;

Alignment Scores:

Pred. No.: 0 Length: 594  
Score: 378.00 Matches: 378  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 33.75% Indels: 0  
DB: 4 Gaps: 0

US-09-270-437D-6 (1-3412) x ABG06795 (1-594)

QY 7 GAGAGGCGGAGGAGCGCGGTACCGGCGGGGGAGCGCGGCTCTCGGGAGAGA 66  
DB 18 GlyGlyGlyGluGluArgArgValProGlyArgGlySerArgGlyLeuSerGlyLysArg 37  
QY 67 CGGATGATGAACAGCTTTACATCGGGAACTGAGCCCGCCCGCTACCGCGGACGACCTC 126  
DB 38 ArgMetMetAsnLysLeuTyrlleGlyAsnLeuSerProAlaValThrAlaAspLeu 57  
QY 127 CGGAGCTCTTTGGGACAGGAAGCTGCCCTCGGGGACAGCTCTCGTGAAGTCCGGC 186  
DB 58 ArgGlnLeuPheGlyAspArgLysLeuProLeuAlaGlyGlnValLeuLysSerGly 77  
QY 187 TAGCCCTTGTGACTACCCGACCGAACTGGGCCATCGCGCCATCGAGACCTCTCG 246  
DB 78 TyrAlaPheValAspTyrProAspGlnAsnTrpAlaIleArgAlaIleGluThrLeuSer 97  
QY 247 GGTAAAGTGAATTGCATGGAAATCATGGAAGTTGATTCTCAGTCTCTAAAGCTA 306  
DB 98 GlyLysValGluLeuHisGlyLysIleMetGluValAspTyrSerValSerLysLysLeu 117  
QY 307 AGGACGAGGAATAATTCAGATTCCGAAATCCCTCTCAGCTGAGGGAGGTGTCGAT 366  
DB 118 ArgSerArgLysIleGlnIleArgAsnIleProProHisLeuGlnTrpGluValLeuAsp 137  
QY 367 GGACTTTGGCTCAATATATGGACAGTGGAGATGTGGAACAAGTCAACACAGACAGAA 426  
DB 138 GlyLeuLeuAlaGlnTyrGlyThrValGluAsnValGluGlnValAsnThrAspThrGlu 157

QY 427 ACCGCGCTTGTCACAGCTCACATATGCAACAAGAGAGCAAAATAGCCATCGAGAAG 486  
DB 158 ThrAlaValValAsnValThrTyraIaThrArgGluGluAlaLysIleAlaMetGluLys 177  
QY 487 CTAAAGCGGCGCATAGTTTGTAGAACTACTCTTCAAGATTTCCTACATCCCGATGAAGAG 546  
DB 178 LeuSerGlyHisGlnPheGluAsnTyrSerPheLysIleSerTyrIleProAspGluGlu 197  
QY 547 GTGAGCTCCCTTCGCGCCCTCAGCGAGCCCGAGCGTGGGACCACTCTCTCCCGGAGCAA 606  
DB 198 ValSerSerProSerProGlnArgAlaGlnArgGlyAspHisSerSerArgGluGln 217  
QY 607 GGCCACGCGCCCTCGGGGCACTTCTCAGCCAGACAGATTGATTTCCGCTCGGATCTCTG 666  
DB 218 GlyHisAlaProGlyGlyThrSerGlnAlaArgGlnIleAspPheProLeuArgIleLeu 237  
QY 667 GTCCCCACCCAGTTTGTGTGGCCATCATCGGAAGAGGCGCTTGACCATAAAGAAACATC 726  
DB 238 ValProThrGlnPheValGlyAlaIleIleGlyLysGluGlyLeuThrIleLysAsnIle 257  
QY 727 ACTAAGCAGACCCAGTCCCGGTAGATATCATAGAAAAGAGAACTCTGGAGCTGCGAG 786  
DB 258 ThrLysGlnThrGlnSerArgValAspIleHisArgLysGluAsnSerGlyAlaAlaGlu 277  
QY 787 AAGCCTCTCACCATCCATGCCACCCAGAGGGGACTTCTGAAGCATGCCCATGATCTT 846  
DB 278 LysProValThrIleHisAlaThrProGluGlyThrSerGluAlaCysArgMetIleLeu 297  
QY 847 GAAATCATGCAAGAAAGGAGGAGATGAGACCAACTAGCCGAAAGAGATTCTCTGAAAATC 906  
DB 298 GluIleMetGlnLysGluAlaAspGluThrLysLeuAlaGluGluIleProLeuLysIle 317  
QY 907 TTGSCACACATGCTGTGGTGGAGACTATTTGGAGACTATTTGAAAATGATATGCTGTTAAC 966  
DB 318 LeuAlaHisAsnGlyLeuValGlyArgLeuIleGlyLysGluGlyArgAsnLeuLysLys 337  
QY 967 ATTGAACATGAAACAGGAGCAAGATAACAATCTCATCTTTGAGGAGTTTGACATATAC 1026  
DB 338 IleGluHisGluThrGlyThrLysIleThrIleSerSerLeuGlnAspLeuSerIleTyr 357  
QY 1027 AACCCGGAAGAACCATCTACTGTGAAGGGCACAGTGTGAGCGCTGTGCGAGTGTGAGATA 1086  
DB 358 AsnProGluArgThrIleThrValLysGlyThrValGluAlaCysAlaSerAlaGluIle 377  
QY 1087 GACATTATGAAGAGCTGCTGAGGCGCTTTGAAAATGATATGCTGCTGTTAAC 1140  
DB 378 GluIleMetLysLysLeuArgGluAlaPheGluAsnAspMetLeuAlaValAsn 395

RESULT 5

ABG96346

ID ABG96346 standard; protein; 556 AA.

XX AC ABG96346;

XX DT 11-DEC-2002 (first entry)

XX DE Human ovarian cancer marker M452.

XX DE Human; ovarian cancer; marker; cancer; familial history; brain disorder;  
KW central nervous system disorder; bacterial meningitis; viral meningitis;  
KW Alzheimer's disease; Parkinson's disease; cerebral oedema; hydrocephalus;  
KW brain herniation; inflammation; encephalitis; testicular disorder;  
KW non-tuberculous granulomatous orchitis; connective tissue disorder;  
KW heart disorder; ischaemic heart disease; atherosclerosis; neoplasm;  
KW histological type; carcinogenic; ovarian cancer marker.

XX OS Homo sapiens.

XX PN WO200271928-A2.

XX PD 19-SEP-2002.

PF 14-MAR-2002; 2002WO-US007826.  
XX  
PR 14-MAR-2001; 2001US-0276025P.  
PR 14-MAR-2001; 2001US-0276026P.  
PR 10-AUG-2001; 2001US-0311732P.  
PR 19-SEP-2001; 2001US-0323580P.  
PR 26-SEP-2001; 2001US-0324967P.  
PR 26-SEP-2001; 2001US-0325102P.  
PR 26-SEP-2001; 2001US-0325149P.  
XX  
PA (MILL-) MILLENNIUM PHARM INC.  
XX  
XX Monahan JE, Gannavarapu M, Hoersch S, Kamatkar S, Kovatis SG;  
PI Meyers RE, Morrisey MP, Olandt PU, Sen A, Vieby PO, Mills GB;  
PI Bast RC, Lu K, Schmandt RE, Zhao X, Glatt K;  
XX  
XX WPI; 2002-723277/78.  
DR N-PSDB; ABS76442.  
XX  
XX  
PT Assessing whether a patient is afflicted with ovarian cancer, useful in  
PT assessing the stage or progression of the disease, comprises comparing  
PT the expression level of a cancer marker in a sample from a patient and  
PT from a non cancer patient.  
XX  
XX  
PS Disclosure; Page 263-264; 48lpp; English.  
XX  
XX The present invention relates to a new method for assessing whether a  
CC patient is afflicted with ovarian cancer. The method involves comparing  
CC the expression level of a marker in a patient sample and the normal level  
CC of expression of the marker in a control non-ovarian cancer sample, where  
CC the marker is selected from 363 cancer markers described in the  
CC specification. The method of the invention is useful in diagnosing or  
CC characterising cancer, in detecting the presence of cancer as early as  
CC possible, and the recurrence of ovarian cancer. The method may also be of  
CC particular use with patients having an enhanced risk of developing  
CC ovarian cancer (e.g. patients having a familial history of ovarian  
CC cancer). The cancer markers may be used in the management and treatment  
CC of e.g. brain and central nervous system disorders (e.g. bacterial and  
CC viral meningitis, Alzheimer's disease or Parkinson's disease), brain  
CC disorders (e.g. cerebral oedema, hydrocephalus or brain herniations),  
CC inflammations (e.g. bacterial or viral meningitis or encephalitis),  
CC testicular disorders (e.g. nonbuberculous granulomatous orchitis),  
CC connective tissue disorders, or heart disorders (e.g. ischaemic heart  
CC disease or atherosclerosis). The compositions and methods may also be  
CC used in assessing the histological type of neoplasm associated with  
CC ovarian cancer, monitoring the progression of ovarian cancer, determining  
CC whether ovarian cancer has metastasized or is likely to metastasize,  
CC selecting a composition for inhibiting ovarian cancer, assessing the  
CC ovarian carcinogenic potential of a compound, or inhibiting ovarian  
CC cancer or at risk of developing ovarian cancer. The present amino acid  
CC sequence represents one of the ovarian cancer markers described in the  
CC invention  
XX  
XX  
SQ Sequence 556 AA;  
  
Alignment Scores:  
Pred. No.: 0 Length: 556  
Score: 357.00 Matches: 357  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 31.87% Indels: 0  
DB: 5 Gaps: 0  
  
US-09-270-437D-6 (1-3412) x ABG96346 (1-556)  
  
QY 70 ATGATGAACAAGCTTTAATCGGGAACCTGAGCCCGCCGCTCACCGCGACGACCTCCGG 129  
Db 1 MetMetAsnLysLeuTyrIleGlyAsnLeuSerProAlaValThrAlaAspLeuArg 20  
QY 130 CAGCTCTTTGGGACAGAGCTGCCCTGGCGGACAGGTCTCTGCTGAAGTCCGGGTAC 189  
Db 21 GlnLeuPheGlyAspArgLysLeuProLeuAlaGlyGlnValLeuLeuLysSerGlyTyr 40

QY 190 GCCTTCGTGGACTACCCGACCCAGAACTGGGCCCATCGCGCCATCGAGACCTCTCCGGT 249  
Db |||||  
41 AlaPheValAspTyrProAspGlnAsnTrpAlaIleArgAlaIleGluThrLeuSerGly 60  
QY 250 AAAGTGAATTCATGCGGAAATAATCATGGAAGTTGATTACTCAGTCTCTTAAAGAGTAAGG 309  
Db |||||  
61 LysValGluLeuHisGlyLysIleMetGluValAspTyrSerValSerLysLysLeuArg 80  
QY 310 AGCAGGAAATTCAGATTCGAAACATCCCTCCTCACCTGCGAGTGGAGGTGTGATGGA 369  
Db |||||  
81 SerArgLysIleGlnIleArgAsnIleProProHisLueGlnTrpGluValLeuAspGly 100  
QY 370 CTTTGGTCAATATGGACAGTGGAGAAATGTGAAACAAGTCAACACACACACACAAACC 429  
Db |||||  
101 LeuLeuAlaGlnTyrGlyThrValGluAsnValGluGlnValAsnThrAspThrGluThr 120  
QY 430 GCCGTGTCAACGTCACATATGCAACAGAGAGAGCAAAATAAGCATCGAGAGAGCTA 489  
Db |||||  
121 AlaValValAsnValThrTyrAlaThrArgGluGluAlaLysIleAlaMetGluLysLeu 140  
QY 490 AGCGGCACTCAGTTTGAGAACTACTCTCTTCAAGATTTCTTACATCCCGGATGAAGAGTG 549  
Db |||||  
141 SerGlyHisGlnPheGluAsnTyrSerPheLysIleSerTyrIleProAspGluGluVal 160  
QY 550 AGCTCCCTTCGCCCTTCAGCGAGCCGAGCGTGGGACCACTCTTCCCGGAGCAAGGC 609  
Db |||||  
161 SerSerProSerProGlnArgAlaGlnArgGlyAspHisSerSerArgGluGlnGly 180  
QY 610 CACGCCCTTGGGGACATTCAGGCGACAGACAGATTGATTTCCCGCTCGGATCTCTGTC 669  
Db |||||  
181 HisAlaProGlyGlyThrSerGlnAlaArgGlnIleAspPheProLeuArgIleLeuVal 200  
QY 670 CCCACCCGATTTGTTGGTGCATCATCGAAAGAGGCGCTTGACCATATAAGAACATCATCT 729  
Db |||||  
201 ProThrGlnPheValGlyAlaIleGlyLysGluGlyLeuThrIleLysAsnIleThr 220  
QY 730 AAGCAGACCCAGTCCCGGTAGATATCCATAGAAAGAGAACTCTGGAGCTCCAGAGAG 789  
Db |||||  
221 LysGlnThrGlnSerArgValAspIleHisArgLysGluAsnSerGlyAlaAlaGluLys 240  
QY 790 CTTGTACCATTCATGCCACCCAGGAGGACTCTGAGCATGCCGATGATCTTGAA 849  
Db |||||  
241 ProValThrIleHisAlaThrProGluGlyThrSerGluAlaCysArgMetIleLeuGlu 260  
QY 850 ATCATGCGAAGACGACAGATGAGACCAACTAGCCGAGAGAGATTCCTCTGAAATCTTG 909  
Db |||||  
261 IleMetGlnLysGluAlaAspGluThrLysLeuAlaGluGluLeuLysIleLeu 280  
QY 910 GCACCAATGGCTTGGTGGAAAGACTGATTGAAAGAGAGGAGAGAAATTTGAAGAAAT 969  
Db |||||  
281 AlaHisAsnGlyLeuValGlyArgLeuIleGlyLysGluGlyArgAsnLeuLysIle 300  
QY 970 GAACATGAACAGGAGCAAGATACAATCTCATCTTTCAGGATTTGAGCATATACAAC 1029  
Db |||||  
301 GluHisGluThrGlyThrLysIleThrIleSerSerLeuGlnAspLeuSerIleTyrAsn 320  
QY 1030 CCGGAAAGAACCATCACTGTGAAGGCGACAGTTGAGGCGCTGTGCCAGTCTCAGATAGAG 1089  
Db |||||  
321 ProGluArgThrIleThrValLysGlyThrValGluAlaCysAlaSerAlaGluIleGlu 340  
QY 1090 ATTATGAAGAGCTGCGTGGCGCTTTGAAATGATATGCTGCTGCTTTAAC 1140  
Db |||||  
341 IleMetLysLysLeuArgGluAlaPheGluAsnAspMetLeuAlaValAsn 357  
  
RESULT 6  
AAUL6166  
ID AAUL6166 standard; protein; 171 AA.  
XX  
AC AAUL6166;  
XX  
DT 07-NOV-2001 (first entry)  
XX Human novel secreted protein, Seq ID 1119.  
DE

XX Human; immunosuppressive; antiarthritic; antirheumatic; cytostatic;  
KW candiant; vasotropic; cerebroprotective; nootropic; neuroprotective;  
KW antibacterial; virucide; fungicide; ophthalmological; vulnery;  
KW secreted protein; rheumatoid arthritis; hyperproliferative disorder;  
KW cardiovascular disorder; cardiac arrest; cerebrovascular disorder;  
KW cerebral ischaemia; angiogenesis; nervous system disorder;  
KW Alzheimer's disease; infection; ocular disorder; corneal infection;  
KW wound healing; epithelial cell proliferation; skin ageing; food additive;  
XX preservative; antiproliferative.  
OS Homo sapiens.  
XX  
XX WO200155322-A2.  
XX  
XX 02-AUG-2001.  
XX  
XX 17-JAN-2001; 2001WO-US001341.  
XX  
XX 31-JAN-2000; 2000US-0179065P.  
PR 04-FEB-2000; 2000US-0180628P.  
PR 24-FEB-2000; 2000US-0184664P.  
PR 02-MAR-2000; 2000US-0186350P.  
PR 16-MAR-2000; 2000US-0189874P.  
PR 17-MAR-2000; 2000US-0190076P.  
PR 18-APR-2000; 2000US-0198123P.  
PR 19-MAY-2000; 2000US-0205151P.  
PR 07-JUN-2000; 2000US-0209467P.  
PR 28-JUN-2000; 2000US-0214886P.  
PR 30-JUN-2000; 2000US-0215135P.  
PR 07-JUL-2000; 2000US-0216647P.  
PR 07-JUL-2000; 2000US-0216880P.  
PR 11-JUL-2000; 2000US-0217487P.  
PR 11-JUL-2000; 2000US-0217496P.  
PR 14-JUL-2000; 2000US-0218290P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 26-JUL-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225113P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225266P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225447P.  
PR 14-AUG-2000; 2000US-0225757P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0226688P.  
PR 22-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 25-SEP-2000; 2000US-0234998P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235836P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 13-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239335P.  
PR 13-OCT-2000; 2000US-0239337P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 01-NOV-2000; 2000US-0241826P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0246474P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249264P.  
PR 17-NOV-2000; 2000US-0249265P.  
PR 17-NOV-2000; 2000US-0249297P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 17-NOV-2000; 2000US-0249300P.  
PR 01-DEC-2000; 2000US-0250160P.  
PR 01-DEC-2000; 2000US-0250391P.  
PR 05-DEC-2000; 2000US-0251030P.  
PR 05-DEC-2000; 2000US-0251988P.



```

PR 27-SEP-2000; 2000US-0235834P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241809P.
PR 01-NOV-2000; 2000US-0244617P.
PR 17-NOV-2000; 2000US-0249299P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
XX
XX (ROSE/) ROSEN C A.
XX (RUBE/) RUBEN S M.
XX (BARA/) BARASH S C.
XX
XX Rosen CA, Ruben SM, Barash SC;
XX
XX WPI; 2003-147444/14.
XX N-PSDB; ABX73494.
XX
XX New polypeptides and nucleic acids, useful in gene therapy for treating,
XX inhibiting or preventing e.g. neural, immune system, muscular,
XX respiratory, reproductive, gastrointestinal, pulmonary, cardiovascular or
XX renal disorders.
XX
XX Claim 11; SEQ ID NO 1119; 402pp; English.
XX
XX The invention relates to human novel polypeptides and their associated
XX polynucleotides. The polypeptides and polynucleotides are useful in gene
XX therapy for treating, inhibiting or preventing neural disorders, immune
XX system disorders (e.g. systemic lupus erythematosus, rheumatoid arthritis
XX and multiple sclerosis), muscular disorders, respiratory diseases (e.g.
XX nasal vestibulitis, nasal polyps and sinusitis), reproductive disorders,
XX gastrointestinal disorders, pulmonary disorders, cardiovascular disorders
XX (e.g. congenital heart defects, Ebstein's anomaly and hypoplastic left
XX heart syndrome), renal disorders (e.g. acute kidney failure and end-stage
XX renal disease), hyperproliferative disorders (e.g. Hodgkin's disease and
XX leukaemia), inflammatory diseases (e.g. septic shock, bursitis and
XX appendicitis), allergic reactions and conditions (e.g. asthma), blood
XX related disorders (e.g. thrombosis, atherosclerosis and myocardial
XX infarction) and cancerous diseases. Sequences ABU54914-ABU55699 and
XX ABU55748 represent human novel polypeptides of the invention
XX
XX SQ Sequence 171 AA;
XX
XX Alignment Scores:
XX Pred. No.: 1e-142 Length: 171
XX Score: 160.00 Matches: 160
XX Percent Similarity: 100.00% Conservative: 0
XX Best Local Similarity: 100.00% Mismatches: 0
XX Query Match: 14.29% Indels: 0
XX DB: 6 Gaps: 0
XX
XX US-09-270-437D-6 (1-3412) x ABU55235 (1-171)
XX
XX 1387 GGGCCATCATCGGGAAGGGGGGACATCAACAGCTGGCAGATTGCGCGAGCC 1446
XX 12 GlyAlaIleIleIleGlyLysGlyAlaHisIleLysGlnLeuAlaArgPheAlaGlyAla 31
XX
XX 1447 TCTATCAGATTCCCGCTCGGGAAGCCCGACGTCAGCGAAGAGTGTTCATCACC 1506
XX 32 SerIleLysIleAlaProAlaGluGlyProAspValSerGluArgMetValIleIleThr 51

```

```

QY 1507 GGGCCACCGGAAGCCAGTTCAAGGCCCCAGGAGCGGATCTTTGGGAAACTGAAAGAGGAA 1566
Db 52 GlyProProGluAlaGlnPheLysAlaGlnGlyArgIlePheGlyLysLeuLysGluGlu 71
QY 1567 AACTTCTTTAAACCCCAAGGAAGTGAAGCTGGAAGCGCATATCAGAGTGCCCTCTTCC 1626
Db 72 AsnPhePheAsnProLysGluGluValLysLeuGluAlaHisIleArgValProSerSer 91
QY 1627 ACAGCTGGCCGGTGAATTGGCAAGGTGGCAAGACCGTGAACGAACCTGCAAGAACTTAACC 1686
Db 92 ThrAlaGlyArgValIleGlyLysGlyIleThrValAsnGluLeuGlnAsnLeuThr 111
QY 1687 AGTGCAGAAGTCATCGTGTCTCTGTCGACCAACCCAGATCAAAATGAGGAAGTGCATGTC 1746
Db 112 SerAlaGluValIleValProArgAspGlnThrProAspGluAsnGluGluValIleVal 131
QY 1747 AGAATTATCGGCACTCTTTGCTAGCCAGACTGCACAGCGCAAGATCAGGGAAATTGTA 1806
Db 132 ArgIleIleGlyHisPhePheAlaSerGlnThrAlaGlnArgLysIleArgGluIleVal 151
QY 1807 CAACAGTGAAGCAGCAGGAGCAGAGAAATACCTCAGGAGTCGCCTCACAGCGCAGCAAG 1866
Db 152 GlnGlnValLysGlnGlnGluGlnLysTyrProGlnGlyValAlaSerGlnArgSerLys 171
XX
XX RESULT 8
XX AAG03261
XX ID AAG03261 standard; protein; 97 AA.
XX AC AAG03261;
XX DT 06-OCT-2000 (first entry)
XX DE Human secreted protein, SEQ ID NO: 7342.
XX KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
XX gene therapy; chromosome mapping.
XX OS Homo sapiens.
XX PN EP1033401-AA.
XX PD 06-SEP-2000.
XX PF 21-FEB-2000; 2000EP-00200610.
XX PR 26-FEB-1999; 99US-0122487P.
XX (GEST ) GENSET.
XX Dumas Milne Edwards J, Duclert A, Giordano J;
XX WPI; 2000-500381/45.
XX N-PSDB; AAC03267.
XX
XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for
XX obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
XX diagnostic, forensic, gene therapy and chromosome mapping procedures.
XX Claim 13; SEQ ID NO 7342; 71pp + Sequence Listing; English.
XX
XX The present sequence is a polypeptide encoded by one of a large number of
XX 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs were
XX prepared from total human RNAs or polyA+ RNAs derived from 30 different
XX tissues. EST sequences usually correspond mainly to the 3' untranslated
XX region (UTR) of the mRNA because they are often obtained from oligo-dT
XX primed cDNA libraries. Such ESTs are not well suited for isolating cDNA
XX sequences derived from the 5' ends of mRNAs and even in those cases where
XX longer cDNA sequences have been obtained, the full 5' UTR is rarely
XX included. 5' ESTs are derived from mRNAs with intact 5' ends and can
XX therefore be used to obtain full length cDNAs and genomic DNAs. 5' ESTs
XX are also used in diagnostic, forensic, gene therapy and chromosome
XX mapping procedures. They are used to obtain upstream regulatory sequences
XX and to design expression and secretion vectors

```

```
XX SQ Sequence 97 AA;
Alignment Scores:
Pred. No.: 1.02e-82 Length: 97
Score: 97.00 Matches: 97
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 8.66% Indels: 0
DB: 3 Gaps: 0

US-09-270-437D-6 (1-3412) x AAG03261 (1-97)
QY 70 ATGATGACAGACITTCATCGGAACTGACCCCGCGCTCAGCCGCGACGACCTCCGG 129
Db 1 MetMetAsnLysLeuTyIleGlyAsnLeuSerProAlaValThrAlaAspAspLeuArg 20
QY 130 CAGCTCTTTGGGGACAGGAAGCTGCCCTGGCGGACAGAGTCCCTGCTGAAGTCCGGCTAC 189
Db 21 GlnLeuPheGlyAspArgLysLeuProLeuAlaGlyGlnValLeuLeuLysSerGlyTyr 40
QY 190 GCCTTCGTGGACTACCCGACAGAACTGGGCGCATCCGCGCATCGAGACCTCTCCGGGT 249
Db 41 AlaPheValAspTyrProAspGlnAsnTrpAlaIleArgAlaIleGluThrLeuSerGly 60
QY 250 AAAGTGGAAATTCAGTGGAAATCATGGAAGTTGATTACTCAGTCTCTAAAAGCTAAGG 309
Db 61 LysValGluLeuHisGlyLysIleMetGluValAspTyrSerValSerLysLysLeuArg 80
QY 310 AGCAGGAAATTCAGATTCGAAACATCCCTCTCCACCTGCAGTGGGAGGTG 360
Db 81 SerArgLysIleGlnIleArgAsnIleProHisLeuGlnTrpGluVal 97

RESULT 9
AAU16164
ID AAU16164 standard; protein; 192 AA.
AC AAU16164;
DT 07-NOV-2001 (first entry)
DE Human novel secreted protein, Seq ID 1117.
KW Human; immunosuppressive; antiarthritic; antirheumatic; cytostatic;
KW cardiant; vasotropic; cerebroprotective; nootropic; neuroprotective;
KW antibacterial; virucide; fungicide; optalmalogical; vulnerary;
KW secreted protein; rheumatoid arthritis; hyperproliferative disorder;
KW cardiovascular disorder; cardiac arrest; cerebrovascular disorder;
KW cerebral ischaemia; angiogenesis; nervous system disorder;
KW Alzheimer's disease; infection; ocular disorder; corneal infection;
KW wound healing; epithelial cell proliferation; skin ageing; food additive;
KW preservative; antiproliferative.
XX OS Homo sapiens.
XX PN WO200155322-A2.
XX XX
XX PD 02-AUG-2001.
XX PF 17-JAN-2001; 2001WO-US001341.
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226686P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 08-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
```



20-OCT-2000; 2000US-0241809P.  
 PR 20-OCT-2000; 2000US-0241826P.  
 PR 01-NOV-2000; 2000US-0244617P.  
 PR 08-NOV-2000; 2000US-0244747P.  
 PR 08-NOV-2000; 2000US-0244757P.  
 PR 08-NOV-2000; 2000US-0244767P.  
 PR 08-NOV-2000; 2000US-0244777P.  
 PR 08-NOV-2000; 2000US-0244787P.  
 PR 08-NOV-2000; 2000US-0245237P.  
 PR 08-NOV-2000; 2000US-0245247P.  
 PR 08-NOV-2000; 2000US-0245257P.  
 PR 08-NOV-2000; 2000US-0245267P.  
 PR 08-NOV-2000; 2000US-0245287P.  
 PR 08-NOV-2000; 2000US-0245337P.  
 PR 08-NOV-2000; 2000US-0246603P.  
 PR 08-NOV-2000; 2000US-0246610P.  
 PR 08-NOV-2000; 2000US-0246611P.  
 PR 08-NOV-2000; 2000US-0246613P.  
 PR 17-NOV-2000; 2000US-0249207P.  
 PR 17-NOV-2000; 2000US-0249208P.  
 PR 17-NOV-2000; 2000US-0249209P.  
 PR 17-NOV-2000; 2000US-0249210P.  
 PR 17-NOV-2000; 2000US-0249211P.  
 PR 17-NOV-2000; 2000US-0249212P.  
 PR 17-NOV-2000; 2000US-0249213P.  
 PR 17-NOV-2000; 2000US-0249214P.  
 PR 17-NOV-2000; 2000US-0249215P.  
 PR 17-NOV-2000; 2000US-0249216P.  
 PR 17-NOV-2000; 2000US-0249217P.  
 PR 17-NOV-2000; 2000US-0249218P.  
 PR 17-NOV-2000; 2000US-0249219P.  
 PR 17-NOV-2000; 2000US-0249245P.  
 PR 17-NOV-2000; 2000US-0249246P.  
 PR 17-NOV-2000; 2000US-0249265P.  
 PR 17-NOV-2000; 2000US-0249297P.  
 PR 17-NOV-2000; 2000US-0249299P.  
 PR 01-DEC-2000; 2000US-0249300P.  
 PR 01-DEC-2000; 2000US-0250160P.  
 PR 01-DEC-2000; 2000US-0250391P.  
 PR 05-DEC-2000; 2000US-0251030P.  
 PR 05-DEC-2000; 2000US-0251988P.  
 PR 05-DEC-2000; 2000US-02556719P.  
 PR 06-DEC-2000; 2000US-0251479P.  
 PR 08-DEC-2000; 2000US-0251856P.  
 PR 08-DEC-2000; 2000US-0251868P.  
 PR 08-DEC-2000; 2000US-0251869P.  
 PR 08-DEC-2000; 2000US-0251989P.  
 PR 11-DEC-2000; 2000US-0254097P.  
 PR 05-JAN-2001; 2001US-02559678P.  
 (HUMA-) HUMAN GENOME SCI INC.  
 Rosen CA, Barash SC, Ruben SM;  
 WPI; 2001-488783/53.  
 N-PSDB; AAS26151.  
 New nucleic acid molecules encoding 461 human secreted proteins for  
 diagnosing, preventing, treating or ameliorating medical conditions and  
 used as food additives or preservatives.  
 Claim 11; SEQ ID NO 1117; 980pp; English.  
 The invention relates to isolated nucleic acid molecules and their  
 encoded secreted proteins. The nucleic acids and proteins are used to  
 prevent, treat or ameliorate a medical condition in e.g. humans, mice,  
 rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used  
 in diagnosing a pathological condition or susceptibility to a  
 pathological condition. Antibodies to the proteins can also be used in  
 alleviating symptoms associated with the disorders and in diagnostic  
 immunoassays e.g. radioimmunoassays or enzyme linked immunosorbant assays

CC (ELISA). Disorders which are diagnosed or treated include autoimmune  
 CC diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g.  
 CC neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac  
 CC arrest, cerebrovascular disorders e.g. cerebral ischaemia, angiogenesis,  
 CC nervous system disorders e.g. Alzheimer's disease, infections caused by  
 CC bacteria, viruses and fungi and ocular disorders e.g. corneal infection,  
 CC and many other disorders listed in the specification. The polypeptides  
 CC can also be used to aid wound healing and epithelial cell proliferation,  
 CC to prevent skin aging due to sunburn, to maintain organs before  
 CC transplantation, for supporting cell culture of primary tissues, to  
 CC regenerate tissues and in chemotaxis. The polypeptides can also be used  
 CC as a food additive or preservative to increase or decrease storage  
 CC capabilities, fat content, lipid, protein, carbohydrate, vitamins,  
 CC minerals, cofactors and other nutritional components. The present,  
 CC sequence represents a novel secreted protein of the invention. Note: The  
 CC sequence data for this patent did not form part of the printed

# Alignment Scores:

Pred. No.: 8.07e-62 Length: 192  
 Score: 75.00 Matches: 75  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 6.70% Indels: 0  
 DB: 4 Gaps: 0

US-09-270-437D-6 (1-3412) x AAU16164 (1-192)

QY 994 ACAATCTCATTTGTCAGGATTTGAGCATATACACCGGAGAAACATCTGTGAAG 1053  
 DB |||||  
 QY 1054 GGCACAGTTGAGGCTGTGCGCAGTGTGAGATAGAGATTATGAGAGCTGCTGAGGCC 1113  
 DB |||||  
 QY 1114 TTGAAATGATATGCTGGCTGTTAACCAAGCCCAATCTGATCCAGGTTGAACCTC 1173  
 DB |||||  
 QY 1174 AGGCACTTGCATCTTTTCAACAGGACTCTCGTGTCTATCTCCA 1218  
 DB |||||  
 132 SerAlaLeuGlyIlePheSerThrGlyLeuSerValLeuSerPro 146

## RESULT 10

ABU55233  
 ID ABU55233 standard; protein; 192 AA.

XX AC ABU55233;

XX DT 18-MAR-2003 (first entry)

XX DE Human novel polypeptide #320.

XX KW Human; neural disorder; immune system disorder; renal disorder;  
 KW muscular disorder; respiratory disease; reproductive disorder;  
 KW gastrointestinal disorder; pulmonary disorder; cardiovascular disorder;  
 KW hyperproliferative disorder; inflammatory disease; allergic reaction;  
 KW blood related disorder; cancer; immunosuppressive; antiinflammatory;  
 KW cardiovascular; nephrotropic; cytostatic; antiallergic; thrombolytic;  
 KW haemostatic; antiarteriosclerotic.

OS Homo sapiens.

XX PN US2002132753-A1.

XX PD 19-SEP-2002.

XX PF 17-JAN-2001; 2001US-00764864.

XX PR 31-JAN-2000; 2000US-0179065P.

PR 04-FEB-2000; 2000US-0180628P.

PR 28-JUN-2000; 2000US-0214886P.

PR 07-JUL-2000; 2000US-0216647P.

PR	07-JUL-2000;	2000US-0216880P.	CC	leukaemia), inflammatory diseases (e.g. septic shock, bursitis and
PR	11-JUL-2000;	2000US-0217487P.	CC	appendicitis), allergic reactions and conditions (e.g. asthma), blood
PR	11-JUL-2000;	2000US-0217496P.	CC	related disorders (e.g. thrombosis, atherosclerosis and myocardial
PR	14-JUL-2000;	2000US-0218290P.	CC	infection) and cancerous diseases. Sequences ABU54914-ABU55699 and
PR	26-JUL-2000;	2000US-0220963P.	CC	ABU55748 represent human novel polypeptides of the invention
PR	26-JUL-2000;	2000US-0220964P.	XX	
PR	14-AUG-2000;	2000US-0224518P.	SQ	Sequence 192 AA;
PR	14-AUG-2000;	2000US-0224519P.		
PR	14-AUG-2000;	2000US-0225267P.		
PR	14-AUG-2000;	2000US-0225268P.		
PR	14-AUG-2000;	2000US-0225270P.		
PR	14-AUG-2000;	2000US-0225447P.		
PR	14-AUG-2000;	2000US-0225757P.		
PR	14-AUG-2000;	2000US-0225758P.		
PR	22-AUG-2000;	2000US-0226868P.		
PR	30-AUG-2000;	2000US-0228924P.		
PR	01-SEP-2000;	2000US-0229287P.		
PR	01-SEP-2000;	2000US-0229343P.		
PR	01-SEP-2000;	2000US-0229344P.		
PR	01-SEP-2000;	2000US-0229345P.		
PR	05-SEP-2000;	2000US-0229509P.		
PR	05-SEP-2000;	2000US-0229513P.		
PR	08-SEP-2000;	2000US-0231413P.		
PR	21-SEP-2000;	2000US-0234223P.		
PR	21-SEP-2000;	2000US-0234274P.		
PR	25-SEP-2000;	2000US-0234997P.		
PR	27-SEP-2000;	2000US-0235834P.		
PR	29-SEP-2000;	2000US-0236327P.		
PR	29-SEP-2000;	2000US-0236367P.		
PR	29-SEP-2000;	2000US-0236368P.		
PR	29-SEP-2000;	2000US-0236369P.		
PR	29-SEP-2000;	2000US-0236370P.		
PR	02-OCT-2000;	2000US-0236802P.		
PR	02-OCT-2000;	2000US-0237037P.		
PR	02-OCT-2000;	2000US-0237038P.		
PR	02-OCT-2000;	2000US-0237039P.		
PR	02-OCT-2000;	2000US-0237040P.		
PR	13-OCT-2000;	2000US-0239935P.		
PR	20-OCT-2000;	2000US-0240960P.		
PR	20-OCT-2000;	2000US-0241785P.		
PR	20-OCT-2000;	2000US-0241809P.		
PR	01-NOV-2000;	2000US-0244617P.		
PR	17-NOV-2000;	2000US-0249299P.		
PR	08-DEC-2000;	2000US-0251856P.		
PR	08-DEC-2000;	2000US-0251868P.		
PR	08-DEC-2000;	2000US-0251869P.		
XX				
PA	(ROSE/) ROSEN C A.			
PA	(RUBE/) RUBEN S M.			
PA	(BARA/) BARASH S C.			
XX				
PI	Rosen CA, Ruben SM, Barash SC;			
XX				
DR	WPI; 2003-147444/14.			
DR	N-PSDB; ABX73492.			
XX				
PT	New polypeptides and nucleic acids, useful in gene therapy for treating,			
PT	inhibiting or preventing e.g. neural, immune system, muscular,			
PT	respiratory, reproductive, gastrointestinal, pulmonary, cardiovascular or			
PT	renal disorders.			
XX				
XX	Claim 11; SEQ ID NO 1117; 402pp; English.			
XX				
CC	The invention relates to human novel polypeptides and their associated			
CC	polynucleotides. The polypeptides and polynucleotides are useful in gene			
CC	therapy for treating, inhibiting or preventing neural disorders, immune			
CC	system disorders (e.g. systemic lupus erythematosus, rheumatoid arthritis			
CC	and multiple sclerosis), muscular disorders, respiratory diseases (e.g.			
CC	nasal vestibulitis, nasal polyps and sinusitis), reproductive disorders,			
CC	gastrointestinal disorders, pulmonary disorders, cardiovascular disorders,			
CC	(e.g. congenital heart defects, Ebstein's anomaly and hypoplastic left			
CC	heart syndrome), renal disorders (e.g. acute kidney failure and end-stage			
CC	renal disease), hyperproliferative disorders (e.g. Hodgkin's disease and			

CC	leukaemia), inflammatory diseases (e.g. septic shock, bursitis and
CC	appendicitis), allergic reactions and conditions (e.g. asthma), blood
CC	related disorders (e.g. thrombosis, atherosclerosis and myocardial
CC	infection) and cancerous diseases. Sequences ABU54914-ABU55699 and
CC	ABU55748 represent human novel polypeptides of the invention
XX	
SQ	Sequence 192 AA;
Alignment Scores:	
Pred. No.:	8.07e-62
Score:	75.00
Percent Similarity:	100.00%
Best Local Similarity:	100.00%
Query Match:	6.70%
DB:	6
US-09-270-437D-6 (1-3412) x ABU55233 (1-192)	
QY	994 ACAATCTCATCTTTGAGGATTTGAGCATATACACCGGAAAGAACCATCTGTGAAG 1053
Db	72 ThrileSerSerLeuGlnAspLeuSerileTyraSnProGluArgThrIleThrVallys 91
QY	1054 GGCACAGTTGAGGCTGTGCCAGTGTGAGATAGATATCAAGAGCTGCGTGAGGCC 1113
Db	92 GlyThrValGluAlaCysAlaSerAlaGluIleGluIleMetLysLysLeuArgGluAla 111
QY	1114 TTTGAAATGATATGCTGCTGTTTAAACCAACAGCAATCTGATCCCGGGTTGAACCTC 1173
Db	112 PheGluAsnAspMetLeuAlaValAsnGlnGlnAlaAsnLeuIleProGlyLeuAsnLeu 131
QY	1174 AGCGCACTTGGCATCTTTTCAACAGGACTGTCCGTGCTATCTTCCA 1218
Db	132 SerAlaLeuGlyIlePheSerThrglyLeuSerValLeuSerPro 146
RESULT 11	
AAU16583	
ID	AAU16583 standard; protein; 171 AA.
XX	
AC	AAU16583;
XX	
DT	07-NOV-2001 (first entry)
DE	Human novel secreted protein, Seq ID 1536.
KW	Human; immunosuppressive; antiarthritic; antirheumatic; cytostatic;
KW	cardiant; vasotropic; cerebroprotective; nootropic; neuroprotective;
KW	antibacterial; virucide; fungicide; ophthalmological; vulnerary;
KW	secreted protein; rheumatoid arthritis; hyperproliferative disorder;
KW	cardiovascular disorder; cardiac arrest; cerebrovascular disorder;
KW	cerebral ischaemia; angiogenesis; nervous system disorder;
KW	Alzheimer's disease; infection; ocular disorder; corneal infection;
KW	wound healing; epithelial cell proliferation; skin ageing; food additive;
XX	preservative; antiproliferative.
OS	Homo sapiens.
XX	
PN	WO200155322-A2.
XX	
PD	02-AUG-2001.
XX	
PF	17-JAN-2001; 2001WO-US001341.
XX	
PR	31-JAN-2000; 2000US-0179065P.
PR	04-FEB-2000; 2000US-0180628P.
PR	24-FEB-2000; 2000US-0184664P.
PR	02-MAR-2000; 2000US-0186350P.
PR	16-MAR-2000; 2000US-0189874P.
PR	17-MAR-2000; 2000US-0190076P.
PR	18-APR-2000; 2000US-0198123P.
PR	19-MAY-2000; 2000US-0205515P.
PR	07-JUN-2000; 2000US-0209467P.
PR	28-JUN-2000; 2000US-0214886P.
PR	30-JUN-2000; 2000US-0215135P.

PR 07-JUL-2000; 2000US-0216647P.  
PR 07-JUL-2000; 2000US-0216880P.  
PR 11-JUL-2000; 2000US-0217487P.  
PR 11-JUL-2000; 2000US-0217496P.  
PR 14-JUL-2000; 2000US-0218290P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 26-JUL-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225213P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225266P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225447P.  
PR 14-AUG-2000; 2000US-0225757P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0226686P.  
PR 22-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 25-SEP-2000; 2000US-0234998P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235836P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239935P.  
PR 13-OCT-2000; 2000US-0239937P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.

PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 20-OCT-2000; 2000US-0241826P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0246474P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249264P.  
PR 17-NOV-2000; 2000US-0249265P.  
PR 17-NOV-2000; 2000US-0249297P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 17-NOV-2000; 2000US-0249300P.  
PR 01-DEC-2000; 2000US-0250160P.  
PR 01-DEC-2000; 2000US-0250391P.  
PR 05-DEC-2000; 2000US-0251030P.  
PR 05-DEC-2000; 2000US-0251988P.  
PR 05-DEC-2000; 2000US-0256719P.  
PR 06-DEC-2000; 2000US-0251479P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.  
PR 08-DEC-2000; 2000US-0251989P.  
PR 11-DEC-2000; 2000US-0251990P.  
PR 11-DEC-2000; 2000US-0254097P.  
PR 05-JAN-2001; 2001US-0259678P.

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Barash SC, Ruben SM;

WPI; 2001-488783/53.

N-PSDB; AAS26570.

PT New nucleic acid molecules encoding 461 human secreted proteins for  
PT diagnosing, preventing, treating or ameliorating medical conditions and  
PT used as food additives or preservatives.

XX Claim 11; SEQ ID NO 1536; 980pp; English.

XX The invention relates to isolated nucleic acid molecules and their  
CC encoded secreted proteins. The nucleic acids and proteins are used to  
CC prevent, treat or ameliorate a medical condition in e.g. humans, mice,  
CC rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used  
CC in diagnosing a pathological condition or susceptibility to a  
CC pathological condition. Antibodies to the proteins can also be used in  
CC alleviating symptoms associated with the disorders and in diagnostic

CC immunoassays e.g. radioimmunoassays or enzyme linked immunosorbant assays  
CC (ELISA). Disorders which are diagnosed or treated include autoimmune  
CC diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g.  
CC neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac  
CC arrest, cerebrovascular disorders e.g. cerebral ischaemia, angiodenesis,  
CC nervous system disorders e.g. Alzheimer's disease, infections caused by  
CC bacteria, viruses and fungi and ocular disorders e.g. corneal infection,  
CC and many other disorders listed in the specification. The polypeptides  
CC can also be used to aid wound healing and epithelial cell proliferation,  
CC to prevent skin aging due to sunburn, to maintain organs before  
CC transplantation, for supporting cell culture of primary tissues, to  
CC regenerate tissues and in chemotaxis. The polypeptides can also be used  
CC as a food additive or preservative to increase or decrease storage  
CC capabilities, fat content, lipid, protein, carbohydrate, vitamins,  
CC minerals, cofactors and other nutritional components. The present  
CC sequence represents a novel secreted protein of the invention. Note: The  
CC sequence data for this patent did not form part of the printed

Alignment Scores:  
Pred. No.: 3.03e-53 Length: 171  
Score: 66.00 Matches: 66  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 5.89% Indels: 0  
DB: 4 Gaps: 0

US-09-270-437D-6 (1-3412) x AAUI6583 (1-171)

QY 1609 ATCAGAGTGCCTCTTCCACAGCTGCGGGTGATTGCAAGTGGCAAGACCGTGAAC 1668  
Db 86 IIEARGVAlProSerSerThrAlaGlyArgValIleGlyLysGlyThrValAsn 105  
QY 1669 GAATCGCAGACTTAACCAAGTCAGAGTCAATCGCTCGTGCAGCAACGCCAGATGAA 1728  
Db 106 GluLeuGlnAsnLeuThrSerAlaGluValIleValProArgAspGlnThrProAspGlu 125  
QY 1729 ATGAGGAGTGCATCGTCAGATTATCGGGCTCTTTGCTAGCCAGACTGCACAGCGC 1788  
Db 126 ASnGluGluValIleValArgIleIleGlyHisPheAlaSerGlnThrAlaGlnArg 145  
QY 1789 AAGATCAGGAAATTTGTA 1806  
Db 146 LysIleArgGluIleVal 151

RESULT 12  
ABU55652  
ID ABU55652 standard; protein; 171 AA.

AC ABU55652;  
XX

DT 18-MAR-2003 (first entry)  
XX

DE Human novel polypeptide #739.  
XX

XX Human: neural disorder; immune system disorder; renal disorder;  
KW muscular disorder; respiratory disease; reproductive disorder;  
KW gastrointestinal disorder; pulmonary disorder; cardiovascular disorder;  
KW hyperproliferative disorder; inflammatory disease; allergic reaction;  
KW blood related disorder; cancer; immunosuppressive; antiinflammatory;  
KW cardiovascular; nephrotropic; cytostatic; antiallergic; thrombolytic;  
KW haemostatic; antiarteriosclerotic.  
XX

OS Homo sapiens.  
XX

XX US2002132753-A1.  
XX

XX 19-SEP-2002.  
XX

XX 17-JAN-2001; 2001US-00764864.  
PF

XX 31-JAN-2000; 2000US-0179065P.  
PR

PR 04-FEB-2000; 2000US-0180628P.  
PR

PR 28-JUN-2000; 2000US-0214886P.  
PR

PR 07-JUL-2000; 2000US-0216647P.  
PR 07-JUL-2000; 2000US-0216880P.  
PR 11-JUL-2000; 2000US-0217487P.  
PR 11-JUL-2000; 2000US-0217496P.  
PR 14-JUL-2000; 2000US-0218290P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 26-JUL-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225447P.  
PR 14-AUG-2000; 2000US-0225757P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 22-AUG-2000; 2000US-0226868P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239935P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.

(ROSE/) ROSEN C A.  
(RUBE/) RUBEN S M.  
(BARA/) BARASH S C.

Rosen CA, Ruben SM, Barash SC;

WPI; 2003-147444/14.

N-PSDB; ABX73911.

XX New polypeptides and nucleic acids, useful in gene therapy for treating,  
PA inhibiting or preventing e.g. neural, immune system, muscular,  
PT respiratory, reproductive, gastrointestinal, pulmonary, cardiovascular or  
PT renal disorders.  
XX

XX Claim 11; SEQ ID NO 1536; 402pp; English.

XX The invention relates to human novel polypeptides and their associated  
CC polynucleotides. The polypeptides and polynucleotides are useful in gene  
CC therapy for treating, inhibiting or preventing neural disorders, immune  
CC system disorders (e.g. systemic lupus erythematosus, rheumatoid arthritis  
CC and multiple sclerosis), muscular disorders, respiratory diseases (e.g.  
CC nasal vestibulitis, nasal polyps and sinusitis), reproductive disorders,  
CC gastrointestinal disorders, pulmonary disorders, cardiovascular disorders  
CC (e.g. congenital heart defects, Ebstein's anomaly and hypoplastic left  
CC heart syndrome), renal disorders (e.g. acute kidney failure and end-stage

CC renal disease), hyperproliferative disorders (e.g. Hodgkin's disease and  
CC leukaemia), inflammatory diseases (e.g. septic shock, bursitis and  
CC appendicitis), allergic reactions and conditions (e.g. asthma), blood  
CC related disorders (e.g. thrombosis, atherosclerosis and myocardial  
CC infarction) and cancerous diseases. Sequences ABU54914-ABU55699 and  
CC ABU55748 represent human novel polypeptides of the invention  
XX  
XX  
SQ Sequence 171 AA;

Alignment Scores:  
Pred. No.: 3.03e-53 Length: 171  
Score: 66.00 Matches: 66  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 5.8% Indels: 0  
DB: 6 Gaps: 0

US-09-270-437D-6 (1-3412) x ABU55652 (1-171)

QY 1609 ATCAGAGTGCCTCTTCCACAGCTGGCGGGTGTATGGCAAGGTGGCAAGACCGTGAAC 1668  
Db |||||||  
QY 1669 GAACGTGAGAACTTAAACAGTGCAGAAATCATCGTCTGCTGACCAACGCCAGATGAA 1728  
Db |||||||  
QY 1729 AATGAGGAAGTATCGTCAGAAATTCATCGGGCACTTCTTTGTACCCAGACTGCACAGCGC 1788  
Db |||||||  
QY 1789 AAGATCAGGGAATTTGA 1806  
Db |||||||

RESULT 13  
ABU89799  
ID ABU89799 standard; protein; 555 AA.  
AC ABU89799;  
DT 10-JUL-2003 (first entry)  
XX Novel human protein NOV14a.  
DE Human; cytostatic; DAPK3-Agonist; DAPK3-Antagonist; cancer; NOV.  
KW Homo sapiens.  
XX  
XX WO2003031571-A2.  
XX  
XX 17-APR-2003.  
XX  
XX 02-OCT-2002; 2002WO-US031357.  
XX

XX 05-OCT-2001; 2001US-0327454P.  
XX 09-OCT-2001; 2001US-0327917P.  
XX 09-OCT-2001; 2001US-0328029P.  
XX 09-OCT-2001; 2001US-0328056P.  
XX 12-OCT-2001; 2001US-0328849P.  
XX 15-OCT-2001; 2001US-0329414P.  
XX 17-OCT-2001; 2001US-0330142P.  
XX 22-OCT-2001; 2001US-0341058P.  
XX 24-OCT-2001; 2001US-0343629P.  
XX 29-OCT-2001; 2001US-0349575P.  
XX 01-NOV-2001; 2001US-0346357P.  
XX 25-JUN-2002; 2002US-0391342P.  
XX 01-OCT-2002; 2002US-00262445.  
XX  
XX (CURA-) CURAGEN CORP.  
XX  
XX

PI Alsobrook JP, Burgess CE, Catterton E, Chant JS, Chaudhuri A;  
PI Edinger SR, Gerlach VL, Giot L, Gorman L, Guo X, Kekuda R;

PI Mezes PS, Millet I, Ooi CE, Patturajan M, Rieger DK, Spytek KA;  
PI Taupier RJ, Zerhusen BD, Zhong H, Zhong M;  
XX  
XX WPI; 2003-381704/36.  
DR N-PSDB; ACA90176.  
XX  
PT New DAPK3 polypeptide, useful for preparing a composition for treating or  
PT preventing e.g., cancer.  
XX  
XX Claim 2; Page 129; 253pp; English.

XX The invention describes an isolated polypeptide comprising any of 33 90-  
CC 1273 amino acid sequences (I) given in the specification or its mature  
CC form, a sequence that is at least 95 % identical to (I), or a sequence  
CC comprising one or more conservative substitutions in the amino acid  
CC sequence of (I). The polypeptide is useful for preparing a composition  
CC for treating or preventing e.g. cancer. This is the amino acid sequence  
CC of a novel human NOV protein  
XX  
SQ Sequence 555 AA;

Alignment Scores:  
Pred. No.: 2.23e-32 Length: 555  
Score: 44.00 Matches: 44  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 3.93% Indels: 0  
DB: 6 Gaps: 0

US-09-270-437D-6 (1-3412) x ABU89799 (1-555)

QY 325 ATTGGAACATCCCTCCTCACCTGCAGTGGAGGTGTGTGATGGACTTTTGGCTCAATAT 384  
Db |||||||  
QY 385 GGGACAGTGGAGAAATGTGGAACAAGTCAACACACACAGAACCCGCGTTGTCAAGTC 444  
Db |||||||  
QY 445 ACATATGCAACA 456  
Db |||||||

RESULT 14  
AAO23971  
ID AAO23971 standard; protein; 187 AA.  
AC AAO23971;  
XX  
XX 06-NOV-2003 (first entry)  
XX  
XX Human IGF-II mRNA binding protein 3.  
XX

XX Anorectic; antiinflammatory; cardiant; hypotensive; antidiabetic;  
KW neuroprotective; pharmaceutical composition; body-weight regulation;  
KW thermogenesis; metabolic; obesity; Syndrome X; insulin-resistance;  
KW eating disorder; cachexia; diabetes mellitus; hypertension; gallstone;  
KW pancreatic dysfunction; arteriosclerosis; coronary heart disease;  
KW hypercholesterolaemia; dyslipidaemia; osteoarthritis; ROS defence;  
KW reactive oxygen species; neurodegenerative; mitochondrial; gene therapy;  
KW human; IGF-II mRNA binding protein 3.  
XX  
XX Homo sapiens.  
XX  
XX WO2003061681-A2.  
XX  
XX 31-JUL-2003.  
XX  
XX 24-JAN-2003; 2003WO-EP000738.  
XX  
XX 25-JAN-2002; 2002EP-00001806.  
XX 14-FEB-2002; 2002EP-00003473.  
XX 28-FEB-2002; 2002EP-00004687.

Thu Jul 22 08:21:41 2004

us-09-270-437d-6.Oligo.rag

PR 25-APR-2002; 2002EP-00009475.  
PR 18-JUN-2002; 2002EP-00013329.  
PR 30-DEC-2002; 2002EP-00029081.  
XX (DEVE-) DEVELOPEN ENTWICKLUNGSBIOLOGISCHE FORSCH.  
PA Steuernagel A, Molitor A, Eulenber K, Broemer G;  
XX WPI; 2003-627418/59.  
XX N-PSDB; AAL57525.  
XX New pharmaceutical composition, useful for the manufacture of an agent  
PT for diagnosing, treating or preventing disorders related to body-weight  
PT regulation and thermogenesis, e.g., metabolic diseases such as obesity.  
XX Claim 3; Fig 7C; 144pp; English.  
XX The invention relates to a novel pharmaceutical composition comprising a  
CC nucleic acid molecule or polypeptide which is a human homologue of a  
CC drosophila melanogaster polypeptide or polynucleotide. The composition of  
CC the invention may be utilised during the diagnosis, study, prevention and  
CC treatment of diseases related to body-weight regulation and thermogenesis  
CC including metabolic disorders such as obesity, Syndrome X and insulin-  
CC resistance syndrome and eating disorders e.g. cachexia, diabetes  
CC mellitus, hypertension, pancreatic dysfunctions, arteriosclerosis,  
CC coronary heart disease, hypercholesterolaemia, dyslipidaemia,  
CC osteoarthritis and gallstones. Furthermore, disorders related to reactive  
CC oxygen species (ROS) defence may be addressed by the invention including  
CC neurodegenerative disorders or mitochondrial disorders. Finally, the  
CC composition of the invention may be useful in gene therapy. The current  
CC sequence is that of the human IGF-II mRNA binding protein 3 protein of  
CC the invention  
XX  
SQ Sequence 187 AA;  
  
Alignment Scores:  
Pred. No.: 6, 92e-21 Length: 187  
Score: 32.00 Matches: 32  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 2, 86% Indels: 0  
DB: 6 Gaps: 0  
  
US-09-270-437D-6 (1-3412) x AAO23971 (1-187)  
QY 1510 CCACCGGAAGCCAGTTCAGGCCAGGACCGATCTTTGGAACTGAAAGAGAAAC 1569  
Db 97 ProProGluAlaGlnPheLysAlaGlnGlyArgIlePheGlyLysLeuLysGluGluAsn 116  
QY 1570 TTCTTTAACCCCAAGAGAGAGTGGAGCGG 1605  
Db 117 PhePheAsnProLysGluGluValLysLeuGluAla 128  
  
RESULT 15  
AAU16581  
ID AAU16581 standard; protein; 47 AA.  
XX  
AC AAU16581;  
XX  
DT 07-NOV-2001 (first entry)  
XX Human novel secreted protein, Seq ID 1534.  
XX Human; immunosuppressive; antiarthritic; antirheumatic; cytostatic;  
KW cardiant; vasotropic; cerebroprotective; nootropic; neuroprotective;  
KW antibacterial; virucide; fungicide; ophthalmological; vulnary;  
KW secreted protein; rheumatoid arthritis; hyperproliferative disorder;  
KW cardiovascular disorder; cardiac arrest; cerebrovascular disorder;  
KW cerebral ischaemia; angiogenesis; nervous system disorder;  
KW Alzheimer's disease; infection; ocular disorder; corneal infection;  
KW wound healing; epithelial cell proliferation; skin ageing; food additive;  
KW preservative; antiproliferative.  
XX

OS Homo sapiens.  
XX WO200155322-A2.  
XX  
XX 02-AUG-2001.  
XX  
XX 17-JAN-2001; 2001WO-US001341.  
XX  
XX 31-JAN-2000; 2000US-0179065P.  
XX 04-FEB-2000; 2000US-0180628P.  
XX 24-FEB-2000; 2000US-0184664P.  
XX 16-MAR-2000; 2000US-0186350P.  
XX 02-MAR-2000; 2000US-0189874P.  
XX 17-MAR-2000; 2000US-0190076P.  
XX 18-APR-2000; 2000US-0198123P.  
XX 19-MAY-2000; 2000US-0205515P.  
XX 07-JUN-2000; 2000US-0209467P.  
XX 28-JUN-2000; 2000US-0214886P.  
XX 30-JUN-2000; 2000US-0215135P.  
XX 07-JUL-2000; 2000US-0216647P.  
XX 07-JUL-2000; 2000US-0216880P.  
XX 11-JUL-2000; 2000US-0217487P.  
XX 11-JUL-2000; 2000US-0217496P.  
XX 14-JUL-2000; 2000US-0218290P.  
XX 26-JUL-2000; 2000US-0220963P.  
XX 26-JUL-2000; 2000US-0220964P.  
XX 14-AUG-2000; 2000US-0224518P.  
XX 14-AUG-2000; 2000US-0224519P.  
XX 14-AUG-2000; 2000US-0225213P.  
XX 14-AUG-2000; 2000US-0225214P.  
XX 14-AUG-2000; 2000US-0225266P.  
XX 14-AUG-2000; 2000US-0225267P.  
XX 14-AUG-2000; 2000US-0225268P.  
XX 14-AUG-2000; 2000US-0225270P.  
XX 14-AUG-2000; 2000US-0225447P.  
XX 14-AUG-2000; 2000US-0225757P.  
XX 14-AUG-2000; 2000US-0225758P.  
XX 18-AUG-2000; 2000US-0225759P.  
XX 22-AUG-2000; 2000US-0226279P.  
XX 22-AUG-2000; 2000US-0226681P.  
XX 22-AUG-2000; 2000US-0226688P.  
XX 22-AUG-2000; 2000US-0227182P.  
XX 23-AUG-2000; 2000US-0227009P.  
XX 30-AUG-2000; 2000US-0228924P.  
XX 01-SEP-2000; 2000US-0229287P.  
XX 01-SEP-2000; 2000US-0229343P.  
XX 01-SEP-2000; 2000US-0229344P.  
XX 05-SEP-2000; 2000US-0229345P.  
XX 05-SEP-2000; 2000US-0229509P.  
XX 06-SEP-2000; 2000US-0229513P.  
XX 06-SEP-2000; 2000US-0230437P.  
XX 06-SEP-2000; 2000US-0230439P.  
XX 08-SEP-2000; 2000US-0231242P.  
XX 08-SEP-2000; 2000US-0231243P.  
XX 08-SEP-2000; 2000US-0231244P.  
XX 08-SEP-2000; 2000US-0231413P.  
XX 08-SEP-2000; 2000US-0231414P.  
XX 08-SEP-2000; 2000US-0232080P.  
XX 08-SEP-2000; 2000US-0232081P.  
XX 12-SEP-2000; 2000US-0231968P.  
XX 14-SEP-2000; 2000US-0232397P.  
XX 14-SEP-2000; 2000US-0232398P.  
XX 14-SEP-2000; 2000US-0232399P.  
XX 14-SEP-2000; 2000US-0232400P.  
XX 14-SEP-2000; 2000US-0232401P.  
XX 14-SEP-2000; 2000US-0233063P.  
XX 14-SEP-2000; 2000US-0233064P.  
XX 21-SEP-2000; 2000US-0233065P.  
XX 21-SEP-2000; 2000US-0234223P.  
XX 21-SEP-2000; 2000US-0234274P.  
XX 25-SEP-2000; 2000US-0234997P.  
XX 25-SEP-2000; 2000US-0234998P.  
XX 26-SEP-2000; 2000US-0235484P.



**This Page Blank (uspto)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: July 16, 2004, 11:16:29 ; Search time 33 Seconds  
(without alignments)  
10675.625 Million cell updates/sec

Title: US-09-270-437D-6  
Perfect score: 1120  
Sequence: 1 ggcaggagagagcgagga.....aaccttgaaaatgtttattt 3412

Scoring table:

Xgapop 60.0 , Xgapext 60.0  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 389414 seqs, 51625971 residues

Word size: 1

Total number of hits satisfying chosen parameters: 663654

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p-model -DEV=xlp

-Q=/cgn2\_1/USTO\_spool\_P/US09270437/runat.16072004.113128.13999/app.query.fasta.1.3591

-DB=issued Patents AA -QFMT=faстан -SUFFIX=Oligo.ra1 -MINMATCH=0.1 -LOPCL=0

-LOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi

-LIST=45 -DOCALIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=15 -MODE=LOCAL

-OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000

-USER=US09270437 @CGN 1.1.46@runat.16072004.113128.13999 -NCFPU=6 -ICPU=3

-NO\_MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSFBLOCK=100 -LONGLOG

-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6

-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

1: Issued Patents AA.\*  
2: /cgn2\_6/ptodata/2/iaa/5A COMB.pcp.\*  
3: /cgn2\_6/ptodata/2/iaa/5B COMB.pcp.\*  
4: /cgn2\_6/ptodata/2/iaa/6A COMB.pcp.\*  
5: /cgn2\_6/ptodata/2/iaa/6B COMB.pcp.\*  
6: /cgn2\_6/ptodata/2/iaa/6C COMB.pcp.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	19	1.7	48	US-09-261-855-20	Sequence 20, Appl
2	19	1.7	577	US-09-261-855-2	Sequence 2, Appl
3	19	1.7	579	US-09-643-597-176	Sequence 176, App
4	19	1.7	579	US-09-643-597-348	Sequence 348, App
5	19	1.7	579	US-09-480-884A-176	Sequence 176, App
6	19	1.7	579	US-09-542-615A-176	Sequence 176, App
7	19	1.7	579	US-09-542-615A-348	Sequence 348, App
8	19	1.7	579	US-09-606-421B-176	Sequence 176, App
9	19	1.7	579	US-09-606-421B-348	Sequence 348, App
10	18	1.6	48	US-09-261-855-24	Sequence 24, Appl
11	16	1.4	49	US-09-261-855-22	Sequence 22, Appl
12	14	1.2	49	US-09-261-855-18	Sequence 18, Appl

13	1.2	47	3	US-09-261-855-21	Sequence 21, Appl
14	1.2	47	3	US-09-261-855-23	Sequence 23, Appl
15	1.0	47	3	US-09-261-855-19	Sequence 19, Appl
16	0.8	30	2	US-08-723-306-25	Sequence 25, Appl
17	0.8	30	2	US-08-723-306-30	Sequence 30, Appl
18	0.8	30	2	PCT-US96-10041-25	Sequence 25, Appl
19	0.8	30	5	PCT-US96-10041-30	Sequence 30, Appl
20	0.8	47	3	US-09-261-855-17	Sequence 17, Appl
21	0.8	81	4	US-09-489-039A-8805	Sequence 8805, Ap
22	0.8	156	4	US-09-252-991A-21912	Sequence 21912, A
23	0.8	399	4	US-09-252-991A-31280	Sequence 31280, A
24	0.8	491	1	US-08-206-176-4	Sequence 4, Appli
25	0.8	603	4	US-09-543-681A-6975	Sequence 6975, Ap
26	0.8	657	4	US-09-252-991A-28001	Sequence 28001, A
27	0.8	863	4	US-09-252-991A-21831	Sequence 21831, A
28	0.7	11	3	US-09-261-855-11	Sequence 11, Appl
29	0.7	11	3	US-09-261-855-13	Sequence 13, Appl
30	0.7	13	3	US-08-602-999A-90	Sequence 90, Appl
31	0.7	13	4	US-09-500-124-90	Sequence 90, Appl
32	0.7	31	3	US-08-602-999A-58	Sequence 58, Appl
33	0.7	31	4	US-09-500-124-58	Sequence 58, Appl
34	0.7	32	4	US-08-278-865-58	Sequence 58, Appl
35	0.7	74	4	US-09-621-976-6160	Sequence 6160, Ap
36	0.7	117	4	US-09-134-001C-5606	Sequence 5606, Ap
37	0.7	136	4	US-09-252-991A-26248	Sequence 26248, A
38	0.7	171	2	US-08-560-098A-53	Sequence 53, Appl
39	0.7	175	4	US-09-252-991A-31050	Sequence 31050, A
40	0.7	204	4	US-09-489-039A-7570	Sequence 7570, Ap
41	0.7	214	4	US-09-489-847-169	Sequence 169, App
42	0.7	228	4	US-09-724-864-42	Sequence 42, Appl
43	0.7	258	4	US-09-252-991A-21821	Sequence 21821, A
44	0.7	258	4	US-09-252-991A-27264	Sequence 27264, A
45	0.7	260	4	US-09-540-236-2747	Sequence 2747, Ap

#### ALIGNMENTS

##### RESULT 1

US-09-261-855-20

; Sequence 20, Application US/09261855A

; Patent No. 6255055

; GENERAL INFORMATION:

; APPLICANT: Ross, Jeffrey

; TITLE OF INVENTION: THE C-MYC CODING REGION DETERMINANT-BINDING PROTEIN

; FILE OF INVENTION: (CRD-BP) AND ITS NUCLEIC ACID SEQUENCE

; CURRENT APPLICATION NUMBER: US/09/261,855A

; CURRENT FILING DATE: 1999-03-03

; NUMBER OF SEQ ID NOS: 46

; SOFTWARE: Patentin Ver. 2.0

; SEQ ID NO 20

; LENGTH: 48

; TYPE: PRT

; ORGANISM: Mus musculus

US-09-261-855-20

Alignment Scores:

Pred. No.:	7.89e-09	Length:	48
Score:	19.00	Matches:	19
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.70%	Indels:	0
DB:	3	Gaps:	0

US-09-270-437D-6 (1-3412) x US-09-261-855-20 (1-48)

QY 1630 GCTGCCGGCTGATTGGCAAGGTCGCAAGCCGTCGAACTCGAAGCTTAACC 1686

Db 8 AlaGlyArgValIleGlyGlyGlyThrValAsnGluLeuGlnAsnLeuThr 26

##### RESULT 2

US-09-261-855-2

; Sequence 2, Application US/09261855A



```
US-09-270-437D-6 (1-3412) x US-09-480-884A-176 (1-579)
QY 1489 AGGATGGTCATCATCCCGGCGCACCGGAGCCAGTTCAGGCCCGAGGACGGATC 1545
Db 452 ArgMetValIleIleThrGlyProGluAlaGlnPheLysAlaGlnGlyArgIle 470

RESULT 6
US-09-542-615A-176
; Sequence 176, Application US/09542615A
; Patent No. 6518256
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy A.
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY
; FILE REFERENCE: 210121.455C8
; CURRENT APPLICATION NUMBER: US/09/542,615A
; CURRENT FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 350
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 176
; LENGTH: 579
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-542-615A-176

Alignment Scores:
Pred. No.: 5,25e-09 Length: 579
Score: 19.00 Matches: 19
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.70% Indels: 0
Db: 4 Gaps: 0

US-09-270-437D-6 (1-3412) x US-09-542-615A-176 (1-579)
QY 1489 AGGATGGTCATCATCCCGGCGCACCGGAGCCAGTTCAGGCCCGAGGACGGATC 1545
Db 452 ArgMetValIleIleThrGlyProGluAlaGlnPheLysAlaGlnGlyArgIle 470

RESULT 7
US-09-542-615A-348
; Sequence 348, Application US/09542615A
; Patent No. 6518256
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy A.
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY
; FILE REFERENCE: 210121.455C8
; CURRENT APPLICATION NUMBER: US/09/542,615A
; CURRENT FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 350
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 348
; LENGTH: 579
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-542-615A-348

Alignment Scores:
Pred. No.: 5,25e-09 Length: 579
Score: 19.00 Matches: 19
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.70% Indels: 0
Db: 4 Gaps: 0

US-09-270-437D-6 (1-3412) x US-09-606-421B-176 (1-579)
QY 1489 AGGATGGTCATCATCCCGGCGCACCGGAGCCAGTTCAGGCCCGAGGACGGATC 1545
Db 452 ArgMetValIleIleThrGlyProGluAlaGlnPheLysAlaGlnGlyArgIle 470

RESULT 8
US-09-606-421B-176
; Sequence 176, Application US/09606421B
; Patent No. 6531315
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy
; APPLICANT: Fanger, Gary R.
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.455C9
; CURRENT APPLICATION NUMBER: US/09/606,421B
; CURRENT FILING DATE: 2000-06-28
; NUMBER OF SEQ ID NOS: 358
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 176
; LENGTH: 579
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-606-421B-176

Alignment Scores:
Pred. No.: 5,25e-09 Length: 579
Score: 19.00 Matches: 19
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.70% Indels: 0
Db: 4 Gaps: 0

US-09-270-437D-6 (1-3412) x US-09-606-421B-176 (1-579)
QY 1489 AGGATGGTCATCATCCCGGCGCACCGGAGCCAGTTCAGGCCCGAGGACGGATC 1545
Db 452 ArgMetValIleIleThrGlyProGluAlaGlnPheLysAlaGlnGlyArgIle 470

RESULT 9
US-09-606-421B-348
; Sequence 348, Application US/09606421B
; Patent No. 6531315
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy
; APPLICANT: Fanger, Gary R.
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.455C9
; CURRENT APPLICATION NUMBER: US/09/606,421B
; CURRENT FILING DATE: 2000-06-28
; NUMBER OF SEQ ID NOS: 358
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 348
```

Score:	16.00	Matches:	16
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.43%	Indels:	0
DB:	3	Gaps:	0
US-09-270-437D-6 (1-3412) x US-09-261-855-22 (1-49)			
QY	925	GTGGAAGACTGATTGGAAAAGAGCGACGAAATTTGAGAAAATTGAA	972
Db	8	ValGlyArgLeuIleGlyGlyGluGlyArgAsnLeuLysIleGlu	23
RESULT 12			
US-09-261-855-18			
; Sequence 18, Application US/09261855A			
; Patent No. 6255055			
; GENERAL INFORMATION:			
; APPLICANT: Ross, Jeffrey			
; TITLE OF INVENTION: THE C-MYC CODING REGION DETERMINANT-BINDING PROTEIN			
; TITLE OF INVENTION: (CRD-BP) AND ITS NUCLEIC ACID SEQUENCE			
; FILE REFERENCE: 960296.95131			
; CURRENT APPLICATION NUMBER: US/09/261,855A			
; CURRENT FILING DATE: 1999-03-03			
; NUMBER OF SEQ ID NOS: 46			
; SOFTWARE: PatentIn Ver. 2.0			
; SEQ ID NO 18			
; TYPE: PRT			
; ORGANISM: Mus musculus			
US-09-261-855-18			
Alignment Scores:			
Pred. No.:	0.000347	Length:	49
Score:	14.00	Matches:	14
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.25%	Indels:	0
DB:	3	Gaps:	0
US-09-270-437D-6 (1-3412) x US-09-261-855-18 (1-49)			
QY	925	GTGGAAGACTGATTGGAAAAGAGCGACGAAATTTGAGAAA	966
Db	8	ValGlyArgLeuIleGlyGlyGluGlyArgAsnLeuLys	21
RESULT 13			
US-09-261-855-21			
; Sequence 21, Application US/09261855A			
; Patent No. 6255055			
; GENERAL INFORMATION:			
; APPLICANT: Ross, Jeffrey			
; TITLE OF INVENTION: THE C-MYC CODING REGION DETERMINANT-BINDING PROTEIN			
; TITLE OF INVENTION: (CRD-BP) AND ITS NUCLEIC ACID SEQUENCE			
; FILE REFERENCE: 960296.95131			
; CURRENT APPLICATION NUMBER: US/09/261,855A			
; CURRENT FILING DATE: 1999-03-03			
; NUMBER OF SEQ ID NOS: 46			
; SOFTWARE: PatentIn Ver. 2.0			
; SEQ ID NO 21			
; TYPE: PRT			
; ORGANISM: Homo sapiens			
US-09-261-855-21			
Alignment Scores:			
Pred. No.:	0.00296	Length:	47
Score:	13.00	Matches:	13
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.16%	Indels:	0
DB:	3	Gaps:	0
US-09-270-437D-6 (1-3412) x US-09-261-855-21 (1-47)			
QY	925	GTGGAAGACTGATTGGAAAAGAGCGACGAAATTTGAGAAA	972
Db	8	ValGlyArgLeuIleGlyGlyGluGlyArgAsnLeuLysIleGlu	23
RESULT 12			
US-09-261-855-18			
; Sequence 18, Application US/09261855A			
; Patent No. 6255055			
; GENERAL INFORMATION:			
; APPLICANT: Ross, Jeffrey			
; TITLE OF INVENTION: THE C-MYC CODING REGION DETERMINANT-BINDING PROTEIN			
; TITLE OF INVENTION: (CRD-BP) AND ITS NUCLEIC ACID SEQUENCE			
; FILE REFERENCE: 960296.95131			
; CURRENT APPLICATION NUMBER: US/09/261,855A			
; CURRENT FILING DATE: 1999-03-03			
; NUMBER OF SEQ ID NOS: 46			
; SOFTWARE: PatentIn Ver. 2.0			
; SEQ ID NO 18			
; TYPE: PRT			
; ORGANISM: Mus musculus			
US-09-261-855-18			
Alignment Scores:			
Pred. No.:	0.000347	Length:	49
Score:	14.00	Matches:	14
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.25%	Indels:	0
DB:	3	Gaps:	0
US-09-270-437D-6 (1-3412) x US-09-261-855-18 (1-49)			
QY	925	GTGGAAGACTGATTGGAAAAGAGCGACGAAATTTGAGAAA	966
Db	8	ValGlyArgLeuIleGlyGlyGluGlyArgAsnLeuLys	21
RESULT 13			
US-09-261-855-21			
; Sequence 21, Application US/09261855A			
; Patent No. 6255055			
; GENERAL INFORMATION:			
; APPLICANT: Ross, Jeffrey			
; TITLE OF INVENTION: THE C-MYC CODING REGION DETERMINANT-BINDING PROTEIN			
; TITLE OF INVENTION: (CRD-BP) AND ITS NUCLEIC ACID SEQUENCE			
; FILE REFERENCE: 960296.95131			
; CURRENT APPLICATION NUMBER: US/09/261,855A			
; CURRENT FILING DATE: 1999-03-03			
; NUMBER OF SEQ ID NOS: 46			
; SOFTWARE: PatentIn Ver. 2.0			
; SEQ ID NO 21			
; TYPE: PRT			
; ORGANISM: Homo sapiens			
US-09-261-855-21			
Alignment Scores:			
Pred. No.:	0.00296	Length:	47
Score:	13.00	Matches:	13
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.16%	Indels:	0
DB:	3	Gaps:	0
US-09-270-437D-6 (1-3412) x US-09-261-855-21 (1-47)			
QY	925	GTGGAAGACTGATTGGAAAAGAGCGACGAAATTTGAGAAA	972
Db	8	ValGlyArgLeuIleGlyGlyGluGlyArgAsnLeuLysIleGlu	23
RESULT 12			
US-09-261-855-18			
; Sequence 18, Application US/09261855A			
; Patent No. 6255055			
; GENERAL INFORMATION:			
; APPLICANT: Ross, Jeffrey			
; TITLE OF INVENTION: THE C-MYC CODING REGION DETERMINANT-BINDING PROTEIN			
; TITLE OF INVENTION: (CRD-BP) AND ITS NUCLEIC ACID SEQUENCE			
; FILE REFERENCE: 960296.95131			
; CURRENT APPLICATION NUMBER: US/09/261,855A			
; CURRENT FILING DATE: 1999-03-03			
; NUMBER OF SEQ ID			

Job time : 49 secs

```
QY      664 CTGGTCCGACCCAGTTTGTGTGTCGATCATCGGAAAG 702
Db      2 LeuValProThrGlnPheValGlyAlaIleIleGlyLys 14

RESULT 14
US-09-261-855-23
; Sequence 23, Application US/09261855A
; Patent No. 6255055
; GENERAL INFORMATION:
; APPLICANT: Ross, Jeffrey
; TITLE OF INVENTION: THE C-MYC CODING REGION DETERMINANT-BINDING PROTEIN
; FILE REFERENCE: 960296.95131
; CURRENT APPLICATION NUMBER: US/09/261,855A
; CURRENT FILING DATE: 1999-03-03
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 23
; LENGTH: 47
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-261-855-23

Alignment Scores:
Pred. No.:      0.00296      Length:      47
Score:          13.00      Matches:      13
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    1.16%      Indels:    0
DB:              3      Gaps:      0

US-09-270-437D-6 (1-3412) x US-09-261-855-23 (1-47)

QY      1432 AGATTGCGGAGCCTCTATCAAGATTGCCCTGCGGAA 1470
Db      24 ArgPheAlaGlyAlaSerIleIleGlyLysIleAlaProAlaGlu 36

RESULT 15
US-09-261-855-19
; Sequence 19, Application US/09261855A
; Patent No. 6255055
; GENERAL INFORMATION:
; APPLICANT: Ross, Jeffrey
; TITLE OF INVENTION: THE C-MYC CODING REGION DETERMINANT-BINDING PROTEIN
; FILE REFERENCE: 960296.95131
; CURRENT APPLICATION NUMBER: US/09/261,855A
; CURRENT FILING DATE: 1999-03-03
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 19
; LENGTH: 47
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-261-855-19

Alignment Scores:
Pred. No.:      0.214      Length:      47
Score:          11.00      Matches:      11
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    0.98%      Indels:    0
DB:              3      Gaps:      0

US-09-270-437D-6 (1-3412) x US-09-261-855-19 (1-47)

QY      1378 CAGGCTGTGGCGCCATCATCGGAGAGAGGGG 1410
Db      6 GlnAlaValGlyAlaIleIleGlyLysGly 16

Search completed: July 16, 2004, 11:30:45
```



```
GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on:      July 16, 2004, 11:29:38 ; Search time 162.5 Seconds
              (without alignments)
              13125.623 Million cell updates/sec

Title:
Perfect score: 1120
Sequence: 1 ggcagcgaggagcgagga.....aacctgaaaatgtttattt 3412

Scoring table:
OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1285345 seqs, 312560633 residues

Word size: 1

Total number of hits satisfying chosen parameters: 2468542

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:
-MODEL=frame+n2p.model -DEV=xlp
-DB=Published Applications AA -QFMT=fastan -SUFFIX=oligo.rapb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo
-TRANS=human40.cdi -LIST=45 -DOCALLIGN=200 -THR SCORE=quality -THR MIN=1
-ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0
-MAXLEN=200000000 -USER=US09270437 @CGN 1 1 250 @runat 16072004 113130 14081
-NCPU=6 -ICPU=3 -NO MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LOGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : Published Applications AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
------------	-------	-------------	--------	----	-------------

1	587	52.4	587	15	US-10-313-986-501	Sequence 501, App
2	511	45.6	620	9	US-09-764-864-1116	Sequence 1116, App
3	357	31.9	556	14	US-10-097-340-147	Sequence 147, App
4	357	31.9	556	16	US-10-648-593-182	Sequence 182, App
5	160	14.3	171	9	US-09-764-864-1119	Sequence 1119, App
6	75	6.7	192	9	US-09-764-864-1117	Sequence 1117, App
7	66	5.9	171	9	US-09-764-864-1536	Sequence 1536, App
8	44	3.9	555	15	US-10-262-445-40	Sequence 40, Appl
9	20	1.8	20	15	US-10-313-986-499	Sequence 499, App
10	20	1.8	47	9	US-09-764-864-1534	Sequence 1534, App
11	19	1.7	48	9	US-09-873-637-20	Sequence 20, Appl
12	19	1.7	81	14	US-10-117-982-476	Sequence 476, App
13	19	1.7	81	15	US-10-313-986-476	Sequence 476, App
14	19	1.7	250	9	US-09-764-864-1532	Sequence 1532, App
15	19	1.7	261	9	US-09-764-864-1114	Sequence 1114, App
16	19	1.7	422	16	US-10-408-765A-2088	Sequence 2088, App
17	19	1.7	577	9	US-09-873-637-2	Sequence 2, Appl
18	19	1.7	577	15	US-10-313-986-500	Sequence 500, App
19	19	1.7	579	9	US-09-735-705-176	Sequence 176, App
20	19	1.7	579	9	US-09-735-705-348	Sequence 348, App
21	19	1.7	579	9	US-09-850-716A-176	Sequence 176, App
22	19	1.7	579	9	US-09-850-716A-348	Sequence 348, App
23	19	1.7	579	9	US-09-897-778-176	Sequence 176, App
24	19	1.7	579	9	US-09-897-778-348	Sequence 348, App
25	19	1.7	579	9	US-09-897-778-446	Sequence 446, App
26	19	1.7	579	9	US-09-897-778-449	Sequence 449, App
27	19	1.7	579	10	US-09-466-396A-176	Sequence 176, App
28	19	1.7	579	12	US-10-007-700-176	Sequence 176, App
29	19	1.7	579	12	US-10-007-700-348	Sequence 348, App
30	19	1.7	579	12	US-10-007-700-446	Sequence 446, App
31	19	1.7	579	12	US-10-007-700-449	Sequence 449, App
32	19	1.7	579	14	US-10-117-982-176	Sequence 176, App
33	19	1.7	579	14	US-10-117-982-348	Sequence 348, App
34	19	1.7	579	14	US-10-117-982-446	Sequence 446, App
35	19	1.7	579	14	US-10-117-982-449	Sequence 449, App
36	19	1.7	579	14	US-10-117-982-480	Sequence 480, App
37	19	1.7	579	14	US-10-117-982-484	Sequence 484, App
38	19	1.7	579	15	US-10-313-986-176	Sequence 176, App
39	19	1.7	579	15	US-10-313-986-348	Sequence 348, App
40	19	1.7	579	15	US-10-313-986-446	Sequence 446, App
41	19	1.7	579	15	US-10-313-986-449	Sequence 449, App
42	19	1.7	579	15	US-10-313-986-480	Sequence 480, App
43	19	1.7	579	15	US-10-313-986-484	Sequence 484, App
44	19	1.7	586	9	US-09-850-716A-427	Sequence 427, App
45	19	1.7	586	9	US-09-897-778-427	Sequence 427, App

## ALIGNMENTS

RESULT 1  
US-10-313-986-501  
; Sequence 501, Application US/10313986  
; Publication No. US20030236209A1  
; GENERAL INFORMATION:  
; APPLICANT: Foy, Teresa M.  
; APPLICANT: McNabb, Andria  
; APPLICANT: Watanabe, Yoshihiro  
; APPLICANT: Reed, Steven G.  
; APPLICANT: Wang, Tonglong  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; CURRENT APPLICATION NUMBER: US/10/313,986  
; FILE REFERENCE: 210121.455C19  
; CURRENT FILING DATE: 2002-12-04  
; NUMBER OF SEQ ID NOS: 560  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 501  
; LENGTH: 587  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-313-986-501  
Alignment Scores:

```

Pred. No.: 0 Length: 587
Score: 587.00 Matches: 587
Percent Similarity: 100.00% Conservative: 0
Best local Similarity: 100.00% Mismatches: 0
Query Match: 52.41% Indels: 0
DB: 15 Gaps: 0

US-09-270-437D-6 (1-3412) x US-10-313-986-501 (1-587)
QY 73 ATGACAAAGCTTTACATCGGAACCTGAGCCCGCCGCTCACCGCCGACGACTCCGGCAG 132
Db 1 MetAsnLysLeuTyrIleGlyAsnLeuSerProAlaValThrAlaAspLeuArgGln 20
QY 133 CTCCTTTGGGACAGAAAGCTGCTCGGGGAGAGTCTCTCAAGTCCGGTACGCC 192
Db 21 LeuPheGlyAspArgLysLeuProLeuAlaGlyGlnValLeuLeuLysSerGlyTyrAla 40
QY 193 TTCGTGGACTACCCCGACACAGAACTGGGCGCATCCGGCCATCGAGACCTCTCCGGTAAA 252
Db 41 PheValAspTyrProAspGlnAsnTrpAlaIleArgAlaIleGluThrLeuSerGlyLys 60
QY 253 GTGGAATTGCATGGGAAATCATGGAAGTTGATTACTCAGTCTCTTAAAGCTAAGGAC 312
Db 61 ValGluLeuHisGlyLysIleMetGluValAspTyrSerValSerLysLysLeuArgSer 80
QY 313 AGGAAATTCAGATTCGAAACATCCCTCTCAGCTGCGAGTGGAGGTGTGGATGACTT 372
Db 81 ArgLysIleGlnIleArgAsnIleProProHisLeuGlnTrpGluValLeuAspGlyLeu 100
QY 373 TTGGCTCAATATGGGACAGTGGAGATGTGGAACAAGTCAACACAGACACAGAAACCGCC 432
Db 101 LeuAlaGlnTyrGlyThrValGluAsnValGluGlnValAsnThrAspThrGluThrAla 120
QY 433 GTTGTCACAGTCATATGCAACAAGAGAGAAAGCAAAATAGCCATGGAGAGCTAAGC 492
Db 121 ValValAsnValTyrTyrAlaThrArgGluGluAlaLysIleAlaMetGluLysLeuSer 140
QY 493 GGGCATCATGTTGAGACTTACTCTTCAAGATTCTCTACATCCCGGATGAGAGGTGAGC 552
Db 141 GlyHisGlnPheGluAsnTyrSerPheLysIleSerTyrIleProAspGluValSer 160
QY 553 TCCCTTCGCCCTCAGAGCCAGCCAGCGTGGGACCACTCTTCCCGGAGCAAGCCAC 612
Db 161 SerProSerProGlnArgAlaGlnArgGlyAspHisSerSerArgGluGlnGlyHis 180
QY 613 GCCCTGGGGGACCTTCTCAGGCCAGACAGATTGATTTCCCGCTCGGATCTGTGTCGCC 672
Db 181 AlaProGlyGlyThrSerGlnAlaArgGlnIleAspPheProLeuArgIleLeuValPro 200
QY 673 AGCCAGTTTGTGGTGCATCATCGAAAGAGGGCTTCACCATAAAGAACATCACTAAG 732
Db 201 ThrGlnPheValGlyAlaIleIleGlyLysGluGlyLeuThrIleLysAsnIleThrLys 220
QY 733 CAGACCCAGTCCCGGTAGATATCCATAGAAAGAGAACTCTGGAGCTGCAGAGAGCCT 792
Db 221 GlnThrGlnSerArgValAspIleHisArgLysGluAsnSerGlyAlaAlaGluLysPro 240
QY 793 GTCACCATCCATGCCACCCAGAGGGGACTTCTGAGCATGCCGATGATTTCTGAATC 852
Db 241 ValThrIleHisAlaThrProGluGlyThrSerGluAlaCysArgMetIleLeuGluIle 260
QY 853 ATGCGAAGAGGACATGAGCAACCACTAGCGGAGAGATTCCTCTGAAAATCTTGCA 912
Db 261 MetGlnLysGluAlaAspGluThrLysLeuAlaGluIleProLeuLysIleLeuAla 280
QY 913 CACAATGGTGTGGTGGAGACTGATTTGAAAGAGAGGAGCAAAATTTGAAGAAATGAA 972
Db 281 HisAsnGlyLeuValGlyArgLeuIleGlyLysGluGlyArgAsnLeuLysIleGlu 300
QY 973 CATGAACAGGGACCAAGATAACAATCTCATCTTTTCAGGATTTGAGCATATACACCG 1032
Db 301 HisGluThrGlyThrLysIleThrIleSerSerLeuGlnAspLeuSerIleTyrAsnPro 320

```

RESULT 2

```

US-09-764-864-1116
; Sequence 1116, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT223
; CURRENT APPLICATION NUMBER: US/09/764,864
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1116
; LENGTH: 620
; TYPE: PRT
; ORGANISM: Homo sapiens

```

```

QY 1033 GAAAGAACCATCACTGTGAAGGCGACAGTTGAGGCGCTGTGCAGTGTGCAGATGAGATT 1092
Db 321 GluArgThrIleThrValLysGlyThrValGluAlaCysAlaSerAlaGluIleGluIle 340
QY 1093 ATGAAGAGCTGCGTGAGGCTTTGAAAATCATATGCTGGCTGTAAACAACAGCCAAAT 1152
Db 341 MetLysLysLeuArgGluAlaPheGluAsnAspMetLeuAlaValAsnGlnAlaAsn 360
QY 1153 CTGATCCAGGGTTGAACCTCAGCGCAGCTTGGCATCTTTCAACAGGAGCTGCGTGCTA 1212
Db 361 LeuIleProGlyLeuAsnLeuSerAlaLeuGlyIlePheSerThrGlyLeuSerValLeu 380
QY 1213 TCTCACCAGCAGGCGCCCGCGAGCTCCCGCGCTACCCCTACACCCCTTCACTACC 1272
Db 381 SerProProAlaGlyProArgGlyAlaProProAlaAlaProTyrHisProPheThr 400
QY 1273 CACTCCGGATATCTTCCAGCTGTACCCCATCACAGTTGGCTTGGCCCGCTTCCGCGAT 1332
Db 401 HisSerGlyTyrPheSerSerLeuTyrProHisGlnPheGlyProPheProHisHis 420
QY 1333 CACTCTTATCCAGAGCAGGAGATTGTGAATCTCTTCATCCCAACCCAGCGTGTGGCGCC 1392
Db 421 HisSerTyrProGluGlnGluIleValAsnLeuPheIleProThrGlnAlaValGlyAla 440
QY 1393 ATCATCGGGAAGAGGGGGCCACACATCAACAGCTGGCGAGATTCCCGGAGCCCTCTATC 1452
Db 441 IleIleGlyLysLysGlyAlaHisIleLysGlnLeuAlaArgPheAlaGlyAlaSerIle 460
QY 1453 AAGATTGCCCTGGGGAAGGCCAGAGCTGACGCGAAAGATGCTCATCATCCCGGCGCA 1512
Db 461 LysIleAlaProAlaGluGlyProAspValSerGluArgMetValIleIleThrGlyPro 480
QY 1513 CCGGAAGCCCATGTTCAAGGCCAGGCGAGTCTTTGGGAACTCTGAAAGAGGAAACTTC 1572
Db 481 ProGluAlaGlnPheLysAlaGlnIleArgIlePheGlyLysLeuLysGluGluAsnPhe 500
QY 1573 TTTAAACCCCAAGAAAGTGAAGCTGGAAGCGCATATCAGAGTGCCTCTTCCACAGCT 1632
Db 501 PheAsnProLysGluGluValLysLeuGluAlaHisIleArgValProSerSerThrAla 520
QY 1633 GCGCGGGTANTGGCAAGGTGGGAAGCCGTAAGACCGTGAACGACTGACAACTTAACCACTGCA 1692
Db 521 GlyArgValIleGlyLysGlyLysThrValAsnGluLeuGlnAsnLeuThrSerAla 540
QY 1693 GAAAGTCATCGTCCCTCGTGACCAAGCCAGATGAAATGAGGAAAGTGCATCGTCAGAAAT 1752
Db 541 GluValIleValProArgAspGlnThrProAspGluAsnGluGluValIleValArgIle 560
QY 1753 ATCGGGCACTTCTTTGTAGCCAGACTGCACAGCGCAAGATCAGGGAATTTGTAACAAG 1812
Db 561 IleGlyHisPhePheAlaSerGlnThrAlaGlnArgLysIleArgGluIleValGlnGln 580
QY 1813 GTGAAGCAGCAGGACAGAAA 1833
Db 581 ValLysGlnGlnGluGlnLys 587

```



```
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (533)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-864-1116

Alignment Scores:
Pred. No.: 0          Length: 620
Score: 511.00        Matches: 511
Percent Similarity: 100.00%  Conservative: 0
Best Local Similarity: 100.00%  Mismatches: 0
Query Match: 45.62%          Indels: 0
DB: 9                      Gaps: 0

US-09-270-437D-6 (1-3412) x US-09-764-864-1116 (1-620)

QY 70 ATGATGAACAAGCTTTACATCGGGAACCTGAGCGCCCGCTGACCCCGACGACCTCCGG 129
Db 22 MetMetAsnLysLeuTyrIleGlyAsnLeuSerProAlaValThrAlaAspLeuArg 41

QY 130 CAGCTCTTTGGGACAGGAAGTCCCTCGCGGGACAGGTCTCTGCTGAAGTCCGGCTAC 189
Db 42 GlnLeuPheGlyAspArgLysLeuProLeuAlaGlyGlnValLeuLeuLysSerGlyTyr 61

QY 190 GCCTTCGTGGACTACCCGACAGAACTGGGCCATCGCGCCATCAGACCCCTCTCGGCT 249
Db 62 AlaPheValAspTyrProAspGlnAsnTrpAlaIleArgAlaIleGluThrLeuSerGly 81

QY 250 AAAGTGGAAATTCATCGGAAATCATGGAAATGATTACTCACTCTCTAAAGCTAAGG 309
Db 82 LysValGluLeuHisGlyLysIleMetGluValAspTyrSerValSerLysLysLeuArg 101

QY 310 AGCAGAAATTCAGATTGAAACATCCCTCTCACCTCGAGTGGAGGTGTTGGATGA 369
Db 102 SerArgLysIleGlnIleArgAsnIleProProHisLeuGlnTrpGluValLeuAspGly 121

QY 370 CTTTGGCTCAATATGGACAGTGGAGATGTGGAACAAGTCAACACAGACACAGAACC 429
Db 122 LeuLeuAlaGlnTyrGlyThrValGluAsnValGluGlnValAsnThrAspThrGluThr 141

QY 430 GCGTGTGTCAAGCTTCATATGCAACAAGAGAGAGCAAAATAGCATGTGGAAGACTA 489
Db 142 AlaValValAsnValThrTyrAlaThrArgGluGluAlaLysIleAlaMetGluLysLeu 161

QY 490 AGCGGCATCAGTTGAGAACTACTCTTCAAGATTCTCTACATCCCGGATGAAGAGTG 549
Db 162 SerGlyHisGlnPheGluAsnTyrSerPheLysIleSerTyrIleProAspGluVal 181

QY 550 AGCTCCCTTCGCGCCCTCAGCGAGCCAGCTGGGGACCACTCTTCCCGGAGCAAGC 609
Db 182 SerSerProSerProProGlnArgAlaGlnArgGlyAspHisSerSerArgGluGlnGly 201

QY 610 CAGCCCTTCGGGGCCTCTCTCAGGCACAGACATTGATTCCCGCTCGCGATCCTGGTC 669
Db 202 HisAlaProGlyGlyThrSerGlnAlaArgGlnIleAspPheProLeuArgIleLeuVal 221

QY 670 CCACCCAGTTTGTGGTCCATCATCGGAAAGAGGGCTTGACCAATAAGAACATCACT 729
Db 222 ProThrGlnPheValGlyAlaIleIleGlyLysGluGlyLeuThrIleLysAsnIleThr 241

QY 730 AAGCAGNCCAGTCCCGGTAGATATCCATAGAAAGAGAACTCTGGAGCTGCAGAGAAG 789
Db 242 LysGlnThrGlnSerArgValAspIleHisArgLysGluAsnSerGlyAlaAlaGluLys 261

QY 790 CTGTCTACCATCCATGCCACCCAGAGGGCACTTCTGAAGCATGCCCATGATTCTTGAA 849
Db 262 ProValThrIleHisAlaThrProGluGlyThrSerGluAlaCysArgMetIleLeuGlu 281

QY 850 ATCATGAGAAAGGAGGATGAGACCAAACTAGCCGAGAGATTCCTCTGAAATCTTG 909
Db 282 IleMetGlnLysGluAlaAspGluThrLysLeuAlaGluIleProLeuLysIleLeu 301

QY 910 GCACACATCGCTTGGTTGGAAGACTGATTGGAAAGAGAGCGCAGAAATTTGAAGAAAT 969
```

```
Db 302 AlaHisAsnGlyLeuValGlyArgLeuIleGlyLysGluGlyArgAsnLeuLysIle 321
QY 970 GAACATGAACAAGGACCAAGATAACAATCTCATCTTTTCAGGATTTGAGCATATACAAC 1029
Db 322 GluHisGluThrGlyThrLysIleThrIleSerSerLeuGlnAspLeuSerIleTyrAsn 341
QY 1030 CCGGAAGAACCACATCACTGTGAAGGCACAGTTGAGCCCTGTGCCAGTGTCTGAGATAGAG 1089
Db 342 ProGluArgThrIleThrValLysGlyThrValGluAlaCysAlaSerAlaGluIleGlu 361
QY 1090 ATTATGAACAAGCTCGCTGAGCCCTTTGAAATGATATCTGCTCTTTAACCAACAGCC 1149
Db 362 IleMetLysLysLeuArgGluAlaPheGluAsnAspMetLeuAlaValAsnGlnAla 381
QY 1150 AATCTCATCCAGGGTTGAACCTCAGCGACCTTGGCATCTTTTCAACAGGACTGTCGGTG 1209
Db 382 AsnLeuIleProGlyLeuAsnLeuSerAlaLeuGlyIlePheSerThrGlyLeuSerVal 401
QY 1210 CTATCTCCACGACAGGCGCCCGGAGCTCCCGCTGCCCCCTTACCAACCCCTTCACT 1269
Db 402 LeuSerProAlaGlyProArgGlyAlaProProAlaAlaProTyrHisProPheThr 421
QY 1270 ACCACTCCGGTACTCTCTCGACCTGTACCCCATCACAGTTTGGCCGCTCCCGCAT 1329
Db 422 ThrHisSerGlyTyrPheSerSerLeuTyrProHisHisGlnPheGlyProPheHis 441
QY 1330 CATCACTCTTATCCAGCAGGAGATTGTGAATCTCTTATCCCAACCCAGGCTGTGGGC 1389
Db 442 HisHisSerTyrProGluGlnGluIleValAsnLeuPheIleProThrGlnAlaValGly 461
QY 1390 GCATCATCGGAAGAAGGGGCGACACATCAACAGCTGGCGAGATTCCCGAGGCTCT 1449
Db 462 AlaIleIleGlyLysGlyAlaHisIleLysGlnLeuAlaArgPheAlaGlyAlaSer 481
QY 1450 ATCAGATTGCCCTCGGAGGCCCGACAGCTCAGCGAAGGATGCTCATCATCAGCGG 1509
Db 482 IleLysIleAlaProAlaGluGlyProAspValSerGluArgMetValIleIleThrGly 501
QY 1510 CCACCGGAAGCCAGTTCAGGCGCCAGGACCGGATCTTTGGGAAACTGAAAGAGGAAAC 1569
Db 502 ProProGluAlaGlnPheLysAlaGlnGlyArgIlePheGlyLysLeuLysGluGluAsn 521
QY 1570 TTCTTTAAACCCAAAGAAAGTGAAGCTGGAA 1602
Db 522 PhePheAsnProLysGluGluValLysLeuGlu 532

RESULT 3
US-10-097-340-147
; Sequence 147, Application US/10097340
; Publication No. US20030087250A1
; GENERAL INFORMATION:
; APPLICANT: John MONAHAN
; APPLICANT: Manjula GANNAVAPU
; APPLICANT: Sebastian HOERSCH
; APPLICANT: Shubhangi KAMATKAR
; APPLICANT: Steve G. KOVATS
; APPLICANT: Rachel E. MEYERS
; APPLICANT: Michael MORRISSEY
; APPLICANT: Peter OLANDT
; APPLICANT: Ami SEN
; APPLICANT: Peter VEIBY
; APPLICANT: Gordon B. MILLS
; APPLICANT: Robert C. BAST, Jr.
; APPLICANT: Karen LU
; APPLICANT: Rosemarie SCHMANDT
; APPLICANT: Xumel ZHAO
; APPLICANT: Karen GLATT
; TITLE OF INVENTION: Nucleic Acid Molecules and Proteins For The Identification,
; TITLE OF INVENTION: Assessment, Prevention, and Therapy of Ovarian Cancer
; FILE REFERENCE: MRI-030
; CURRENT APPLICATION NUMBER: US/10/097,340
; CURRENT FILING DATE: 2002-03-14
```

;  
; PRIOR APPLICATION NUMBER: 60/276,025  
; PRIOR FILING DATE: 2001-03-14  
; PRIOR APPLICATION NUMBER: 60/325,149  
; PRIOR FILING DATE: 2001-09-26  
; PRIOR APPLICATION NUMBER: 60/276,026  
; PRIOR FILING DATE: 2001-03-14  
; PRIOR APPLICATION NUMBER: 60/324,967  
; PRIOR FILING DATE: 2001/09/26  
; PRIOR APPLICATION NUMBER: 60/311,732  
; PRIOR FILING DATE: 2001-08-10  
; PRIOR APPLICATION NUMBER: 60/325,102  
; PRIOR FILING DATE: 2001-09-26  
; PRIOR APPLICATION NUMBER: 60/323,580  
; PRIOR FILING DATE: 2001-09-19  
; NUMBER OF SEQ ID NOS: 363  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 147  
; LENGTH: 556  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-097-340-147

Alignment Scores:  
Pred. No.: 9,24e-315 Length: 556  
Score: 357.00 Matches: 357  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 31.87% Indels: 0  
DB: 14 Gaps: 0

US-09-270-437D-6 (1-3412) x US-10-097-340-147 (1-556)

QY 70 ATGATGACAAAGCTTACATCGGAACTGAGCCCGCCGCTACCGCGAGCACTCCGG 129  
Db 1 MetMetAsnLysLeuThrIleGlyAsnLeuSerProAlaValThrAlaAspLeuArg 20  
QY 130 CAGCTCTTTGGGGACAGAAAGCTGCTGGGGACAGAGTCTGCTGAAGTCCGGCTAC 189  
Db 21 GlnLeuPheGlyAspArgLysLeuProLeuAlaGlyGlnValLeuLeuLysSerGlyTyr 40  
QY 190 GCCTTCGTGGACTACCGGACAGAACTGGGGCCATCCCGCCGATCGAGACCTCTCGGGT 249  
Db 41 AlaPheValAspTyrProAspGlnAsnTrpAlaIleArgAlaIleGluThrLeuSerGly 60  
QY 250 AAAGTGGAAATTCAGATTCGAAATCATGGAAGTTGATTACTCAGTCTCTAAAAGCTAAGG 309  
Db 61 LysValGluLeuHisGlyLysIleMetGluValAspTyrSerValSerLysLeuArg 80  
QY 310 AGCAGAAATTCAGATTCGAAATCATCCCTCTCAGTGCAGTGGAGGTGTGGATGGA 369  
Db 81 SerArgLysIleGlnIleArgAsnIleProHisLeuGlnTrpGluValLeuAspGly 100  
QY 370 CTTTGGCTCAATATGGACAGTGGAGATGTGGNACAGTCAAGTCAACACAGACAGAAACC 429  
Db 101 LeuLeuAlaGlnTyrGlyThrValGluAsnValGluGlnValAsnThrAspThrGluThr 120  
QY 430 GCCCTTGTCAACGTCACATATGCAACAGAGAAGCAAAATAGCCATCGAGAAGCTA 489  
Db 121 AlaValValAsnValThrTyrAlaThrArgGluGluAlaLysIleAlaMetGluLysLeu 140  
QY 490 AGCGGACATCAGTTTGAAACTACTCTTCAAGATTTCTCATATCCCGGATCAAGAGGTG 549  
Db 141 SerGlyHisGlnPheGluAsnTyrSerPheLysIleSerTyrIleProAspGluGluVal 160  
QY 550 AGCTCCCTTCGCCCTCAGGAGCCAGCGTGGGACCACTCTTCCCGGAGCAGGC 609  
Db 161 SerSerProSerProGlnHarGalaGlnArgGlyAspHisSerSerArgGluGlnGly 180  
QY 610 CACGCCCTTGGGGCACTTCTCAGGCCACAGATTGATTTCCCGCTCGGATCCTGGTC 669  
Db 181 HisAlaProGlyGlyThrSerGlnAlaArgGlnIleAspPheProLeuArgIleLeuVal 200  
QY 670 CCCACCAAGTTTGTGGTGCATCATCGGAAAGGAGGCTTGACCATAAAGAAATCACT 729

Db 201 ProThrGlnPheValGlyAlaIleIleGlyLysGluGlyLeuThrIleLysAsnIleThr 220  
QY 730 AAGCAGACCCAGTCCCGGTAGATATCCATAGAAAGAGAACTCTGGAGCTGCAGAGAAG 789  
Db 221 LysGlnThrGlnSerArgValAspIleHisArgLysGluAsnSerGlyAlaAlaGluLys 240  
QY 790 CCTGTCAACCATCCATGCCACCCAGAGGGACTTCTGAGCATGCCGATATCTTGAA 849  
Db 241 ProValThrIleHisAlaThrProGluGlyThrSerGluAlaCysArgMetIleLeuGlu 260  
QY 850 ATCATGCAAAAGAGGAGATGACCAAACTAGCCCAAGAGATTCTCTCTGAAATCTTG 909  
Db 261 IleMetGlnLysGluAlaAspGluThrLysLeuAlaGluGluIleProLeuLysIleLeu 280  
QY 910 GCACACAATGGCTTGGTGAAGACTGATTGGAAGAAAGCAGAAATTTGAAGAAAT 969  
Db 281 AlaHisAsnGlyLeuValGlyArgLeuIleGlyLysGluGlyArgAsnLeuLysLysIle 300  
QY 970 GAACATGAAACAGGAGCCAGATAAATCTCATCTTTGACAGGATTTGAGCATATACAAC 1029  
Db 301 GluHisGluThrGlyThrLysIleThrIleSerSerLeuGlnAspLeuSerIleTyrAsn 320  
QY 1030 CCGGAAAGAACCATCACTGTGAAGGACACAGTTCAGGCCTGTGCAGTCTGAGATAGAG 1089  
Db 321 ProGluArgThrIleThrValLysGlyThrValGluAlaCysAlaSerAlaGluIleGlu 340  
QY 1090 ATTATGAAGAGCTGCGTGAGGCCTTTGAAAATGATATGCTGGCTGTTTAAAC 1140  
Db 341 IleMetLysLysLeuArgGluAlaPheGluAsnAspMetLeuAlaValAsn 357

RESULT 4

US-10-648-593-182  
; Sequence 182, Application US/10648593  
; Publication No. US20040106132A1  
; GENERAL INFORMATION:  
; APPLICANT: Bristol-Myers Squibb Company  
; TITLE OF INVENTION: IDENTIFICATION OF GENES FOR PREDICTING ACTIVITY OF COMPOUNDS THAT  
; TITLE OF INVENTION: INTERACT WITH AND/OR MODULATE PROTEIN TYROSINE KINASES AND/OR  
; TITLE OF INVENTION: PROTEIN TYROSINE KINASE PATHWAYS IN BREAST CELLS  
; FILE REFERENCE: D0273 NP  
; CURRENT APPLICATION NUMBER: US/10/648,593  
; CURRENT FILING DATE: 2003-08-26  
; PRIOR APPLICATION NUMBER: 60/406,385  
; PRIOR FILING DATE: 2002-08-27  
; NUMBER OF SEQ ID NOS: 557  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 182  
; LENGTH: 556  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-648-593-182

Alignment Scores:  
Pred. No.: 9,24e-315 Length: 556  
Score: 357.00 Matches: 357  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 31.87% Indels: 0  
DB: 16 Gaps: 0

US-09-270-437D-6 (1-3412) x US-10-648-593-182 (1-556)

QY 70 ATGATGACAAAGCTTACATCGGAACTGAGCCCGCCGCTACCGCGAGCACTCCGG 129  
Db 1 MetMetAsnLysLeuThrIleGlyAsnLeuSerProAlaValThrAlaAspLeuArg 20  
QY 130 CAGCTCTTTGGGGACAGAAAGCTGCTGGGGACAGAGTCTCTCTGCAAGTCCGGCTAC 189  
Db 21 GlnLeuPheGlyAspArgLysLeuProLeuAlaGlyGlnValLeuLeuLysSerGlyTyr 40  
QY 190 GCCTTCGTGGACTACCGGACAGAACTGGGGCCATCCCGCCGATCGAGACCTCTCGGGT 249

```
Db 41 AlaPheValAspTyrProAspGlnAsnTrpAlaIleAlaIleGluThrIleuSerGly 60
QY 250 AAAGTGGAAATTCATCGGAAATCATGAAGTTGATTACTAGTCTCTAAAGCTAAGG 309
Db 61 LysValGluLeuHisGlyLeuMetGluValAspTyrSerValSerLysLysLeuArg 80
QY 310 AGCAGGAAATTCAGATTTCGAACATCCCTCTCCTACCTGCGAGTGGAGGTGTGGATGGA 369
Db 81 SerArgLysIleGlnIleArgAsnIleProProHisLeuGlnTrpGluValIleuAspGly 100
QY 370 CTTTGGCTCAATATGGACAGTGGAGAAATGTGGAACAAAGTCAACACAGACACAGAAACC 429
Db 101 LeuLeuAlaGlnTyrGlyThrValGluAsnValGluGlnValAsnThrAspThrGluThr 120
QY 430 GCGCTGTTCACAGTCACATATGCCACAGACAGAGAGCAAAAAATAGCCATCGAGAGCTA 489
Db 121 AlaValValAsnValThrTyrAlaThrArgGluGluAlaLysIleAlaMetGluLysLeu 140
QY 490 AGCGGGCATCAGTTTGGAACTACTCTTCAAGATTTCTACATCCCGGATGAGAGGTG 549
Db 141 SerGlyHisGlnPheGluAsnTyrSerPheLysIleSerTyrIleProAspGluGluVal 160
QY 550 AGCTCCCTTCGCCCTCAGCAGCCAGCGTGGGACCACTCTTCCCGGGAGCAAGGC 609
Db 161 SerSerProSerProProGlnArgAlaGlnArgGlyAspHisSerSerArgGluGlnGly 180
QY 610 CACGCCCTGGGGCACTCTCAGGCCAGACAGATTGATTTCCGCTCGGATCCTCGTC 669
Db 181 HisAlaProGlyGlyThrSerGlnAlaArgGlnIleAspPheProLeuArgIleLeuVal 200
QY 670 CCCACCCAGTTTGTGTGGTCCATCATCGGAAGAGGCTTGACCATPAAAGACATCACT 729
Db 201 ProThrGlnPheValGlyAlaIleIleGlyLysGluGlyLeuThrIleLysAsnIleThr 220
QY 730 AAGCAGACCCAGTCCCGGTAGATATCCATAGAAAGAGAACTCTGGAGTGCAGAGAAG 789
Db 221 LysGlnThrGlnSerArgValAspIleHisArgLysGluAsnSerGlyAlaAlaGluLys 240
QY 790 CTTGTACCATCATCCATCCACCCAGAGGACTTCTGAAGCATGCCCATGATTCTTGAA 849
Db 241 ProValThrIleHisAlaThrProGluGlyThrSerGluAlaCysArgMetIleLeuGlu 260
QY 850 ATCATCAGAAAGAGGAGATGAGACCAACTAGCCGAAGAGATTCCTCTGAAATCTTG 909
Db 261 IleMetGlnLysGluAlaAspGluThrLysLeuAlaGluGluLeuProLeuLysIleLeu 280
QY 910 GCACACAATGGCTGTGGTGAAGCTGATTGGAAGAGAGCCAGAAATTTGAAGAAAATT 969
Db 281 AlaHisAsnGlyLeuValGlyArgLeuIleGlyLysGluGlyArgAsnLeuLysLysIle 300
QY 970 GAACATGAACAGGACGAGATAACAATCTCTCTTTCAGGATTTGACCATATCAAC 1029
Db 301 GluHisGluThrGlyThrLysIleThrIleSerSerLeuGlnAspLeuSerIleTyrAsn 320
QY 1030 CCGGAAGAACCATCACTGTGAAGGCGACAGTTGAGCCCTGTGCCAGTGTCTGAGATAG 1089
Db 321 ProGluArgThrIleThrValLysGlyThrValGluAlaCysLysAlaSerAlaGluIleGlu 340
QY 1090 ATTATCAAGAGCTGGTGGGCTTTGAAATGATATGATGCTGCTTAAAC 1140
Db 341 IleMetLysLysLeuArgGluAlaPheGluAsnAspMetLeuAlaValAsn 357
```

RESULT 5

```
US-09-764-864-1119
; Sequence 1119, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT23
; CURRENT APPLICATION NUMBER: US/09/764,864
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
```

```
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1119
; LENGTH: 171
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-764-864-1119
```

Alignment Scores:

```
Pred. No.: 8,27e-136 Length: 171
Score: 160.00 Matches: 160
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 14,29% Indels: 0
DB: 9 Gaps: 0
```

US-09-270-437D-6 (1-3412) x US-09-764-864-1119 (1-171)

```
QY 1387 GCGCCATCATCGGAAAGAGGGGCGACACATCAACAGCTGGCGAGATTCGCCGAGCC 1446
Db 12 GlyAlaIleIleGlyLysGlyAlaHisIleLysGlnLeuAlaArgPheAlaGlyAla 31
QY 1447 TCTATCAAGATTGCCCTCGGAAAGGCCCGACAGCTCAGCGAAAGATGGTCATCATCACC 1506
Db 32 SerIleLysIleAlaProAlaGluGlyProAspValSerGluArgMetValIleIleThr 51
QY 1507 GGGCCACCGGAAGCCCAAGTTCAAGGCCCGAGGCGGATCTTGGGAACTGAAAGAGAA 1566
Db 52 GlyProGluAlaGlnPheLysAlaGlnGlyArgIlePheGlyLysLeuLysGluGlu 71
QY 1567 AACTTCTTTAAACCCCAAGAAAGTGAAGCTGGAAGCGCATATCAGAGTGCCTCTTCC 1626
Db 72 AsnPheAsnProLysGluGluValLysLeuGluAlaHisIleArgValProSerSer 91
QY 1627 ACAGCTGGCCGGGTGATTGCCAAAGGTGGCAAGACCGTGAACCACTGCAAGACTTAACC 1686
Db 92 ThrAlaGlyArgValIleGlyLysGlyLysThrValAsnGluLeuGlnAsnLeuThr 111
QY 1687 AGTGCAGAACTCATCGTCTGTGACCAACCCAGATGAAATGAGGAAGTATCGTC 1746
Db 112 SerAlaGluValIleValProArgAspGlnThrProAspGluAsnGluGluValIleVal 131
QY 1747 AGAATATCGGCCTCTTCTTGTAGCAGACTGCACAGCGCAAGATCAGGAAATGTA 1806
Db 132 ArgIleIleGlyHisPheAlaSerGlnThrAlaGlnArgLysIleArgGluIleVal 151
QY 1807 CAACAGGTGAACGACGAGGAGCAAGAAATACCTCAGGAGTCCCTCAGCGCAGCAAG 1866
Db 152 GlnGlnValLysGlnGlnGluGlnLysTyrProGlnGlyValAlaSerGlnArgSerLys 171
```

RESULT 6

```
US-09-764-864-1117
; Sequence 1117, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT23
; CURRENT APPLICATION NUMBER: US/09/764,864
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1117
; LENGTH: 192
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (71)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (147)
```

OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
NAME/KEY: SITE  
LOCATION: (160)  
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
NAME/KEY: SITE  
LOCATION: (164)  
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
NAME/KEY: SITE  
LOCATION: (165)  
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
NAME/KEY: SITE  
LOCATION: (168)  
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
US-09-764-864-1117

Alignment Scores:  
Pred. No.: 1.3e-58 Length: 192  
Score: 75.00 Matches: 75  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 6.70% Indels: 0  
DB: 9 Gaps: 0

US-09-270-437D-6 (1-3412) x US-09-764-864-1117 (1-192)

QY 994 ACAATCTCATCTTTGAGGATTGAGCATATACAAACCGGAAGAACCATCTCTGTGAAG 1053  
Db 72 ThrilleSerSerLeuGlnAspLeuSerIleTyAsnProGluArgThrIleThrVallys 91  
QY 1054 GCACAGTGTAGGCTGTGCCAGTGTGAGATAGAGATTATGAAGAAGCTGCTGAGGCC 1113  
Db 92 GlyThrValGluAlaCysAlaSerAlaGluIleGluIleMeCysLysLeuArgGluAla 111  
QY 1114 TTTGAAATGATATGTGCTGCTTTAACCAACAGCAATCTGATCCAGGGTTGAACCTC 1173  
Db 112 PheGluAsnAspMetLeuAlaValAsnGlnAlaAsnLeuIleProGlyLeuAsnLeu 131  
QY 1174 ACGGACTGGCATCTTTTCAACAGACTGTCCGGTGCATCTCCA 1218  
Db 132 SerAlaLeuGlyIlePheSerThrGlyLeuSerValLeuSerPro 146

RESULT 7  
US-09-764-864-1536  
; Sequence 1536, Application US/09764864  
; Patent No. US20020132753A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: FTZ23  
; CURRENT APPLICATION NUMBER: US/09/764,864  
; CURRENT FILING DATE: 2001-01-17  
; Prior application data removed - consult PALM or file wrapper  
; NUMBER OF SEQ ID NOS: 1792  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1536  
; LENGTH: 171  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SITE  
; LOCATION: (65)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
; NAME/KEY: SITE  
; LOCATION: (170)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
US-09-764-864-1536

Alignment Scores:  
Pred. No.: 1.98e-50 Length: 171  
Score: 66.00 Matches: 66  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 5.89% Indels: 0

DB: 9 Gaps: 0  
US-09-270-437D-6 (1-3412) x US-09-764-864-1536 (1-171)  
QY 1609 ATCAGAGTGCCTCTTCCACAGCTGGCGGGTGATTGGCAAGGTGGCAAGACCGTGAAC 1668  
Db 86 IleArgValProSerSerThrAlaGlyArgValIleGlyLysThrValAsn 105  
QY 1669 GAATCGAGAACTTAACAGTGCAGAACTCATCTGCTCGTGCACCAACCCAGATCAA 1728  
Db 106 GluLeuGlnAsnLeuThrSerAlaGluValIleValProArgAspGlnThrProAspGlu 125  
QY 1729 AATCAGGAAGTATCGTCAGAAATATATCGGCACACTCTTTTGTAGCCAGACTGCACAGCGC 1788  
Db 126 AsnGluGluValIleValArgIleIleGlyHisPheAlaSerGlnThrAlaGlnArg 145  
QY 1789 AAGATCAGGGAATTTGTA 1806  
Db 146 LysIleArgGluIleVal 151  
RESULT 8  
US-10-262-445-40  
; Sequence 40, Application US/10262445  
; Publication No. US20040014058A1  
; GENERAL INFORMATION:  
; APPLICANT: Alsobrook II, John  
; APPLICANT: Burgess, Catherine  
; APPLICANT: Catterton, Elina  
; APPLICANT: Chant, John  
; APPLICANT: Chaudhuri, Amitabha  
; APPLICANT: Edinger, Shlomit  
; APPLICANT: Gerlach, Valerie  
; APPLICANT: Giot, Loic  
; APPLICANT: Gorman, Linda  
; APPLICANT: Guo, Xiaojia  
; APPLICANT: Kekuda, Ramesh  
; APPLICANT: Mezes, Peter  
; APPLICANT: Millet, Isabelle  
; APPLICANT: Ooi, Chean Eng  
; APPLICANT: Patturajan, Meera  
; APPLICANT: Rieger, Daniel  
; APPLICANT: Spytek, Kimberly  
; APPLICANT: Taupier Jr., Raymond J.  
; APPLICANT: Zerhusen, Bryan  
; APPLICANT: Zhong, Haihong  
; APPLICANT: Zhong, Mei  
; TITLE OF INVENTION: NOVEL HUMAN PROTEINS, POLYNUCLEOTIDES ENCODING THEM AND METHODS  
; FILE REFERENCE: 21402-462D  
; CURRENT APPLICATION NUMBER: US/10/262,445  
; CURRENT FILING DATE: 2002-10-01  
; PRIOR APPLICATION NUMBER: 60/327,454  
; PRIOR FILING DATE: 2001-10-05  
; PRIOR APPLICATION NUMBER: 60/327,917  
; PRIOR FILING DATE: 2001-10-09  
; PRIOR APPLICATION NUMBER: 60/328,029  
; PRIOR FILING DATE: 2001-10-09  
; PRIOR APPLICATION NUMBER: 60/328,056  
; PRIOR FILING DATE: 2001-10-09  
; PRIOR APPLICATION NUMBER: 60/328,849  
; PRIOR FILING DATE: 2001-10-12  
; PRIOR APPLICATION NUMBER: 60/329,414  
; PRIOR FILING DATE: 2001-10-15  
; PRIOR APPLICATION NUMBER: 60/330,142  
; PRIOR FILING DATE: 2001-10-17  
; PRIOR APPLICATION NUMBER: 60/341,058  
; PRIOR FILING DATE: 2001-10-22  
; PRIOR APPLICATION NUMBER: 60/343,629  
; PRIOR FILING DATE: 2001-10-24  
; PRIOR APPLICATION NUMBER: 60/349,575  
; PRIOR FILING DATE: 2001-10-29  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 133



```
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Mericle, Barbara
; APPLICANT: Spies, Gregory A.
; APPLICANT: Fan, Liqun
; APPLICANT: Wang, Tongtong
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C18
; CURRENT APPLICATION NUMBER: US/10/117,982
; CURRENT FILING DATE: 2002-04-05
; NUMBER OF SEQ ID NOS: 484
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 476
; LENGTH: 81
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-117-982-476

Alignment Scores:
Pred. No.: 118-07 Length: 81
Score: 19.00 Matches: 19
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.70% Indels: 0
DB: 14 Gaps: 0

US-09-270-437D-6 (1-3412) x US-10-117-982-476 (1-81)
Qy 1489 AGGATGTCATCATCACCGGCCACCGAAGCCAGTTCAGGCCCGGACGGATC 1545
Db 24 ArgMetValIleIleThrGlyProGluAlaGlnPheLysAlaGlnGlyArgIle 42

RESULT 13
US-10-313-986-476
; Sequence 476, Application US/10313986
; Publication No. US20030236209A1
; GENERAL INFORMATION:
; APPLICANT: Poy, Teresa M.
; APPLICANT: McNabb, Andria
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Reed, Steven G.
; APPLICANT: Wang, Tongtong
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C19
; CURRENT APPLICATION NUMBER: US/10/313,986
; CURRENT FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 560
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 476
; LENGTH: 81
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-313-986-476

Alignment Scores:
Pred. No.: 118-07 Length: 81
Score: 19.00 Matches: 19
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.70% Indels: 0
DB: 15 Gaps: 0

US-09-270-437D-6 (1-3412) x US-10-313-986-476 (1-81)
Qy 1489 AGGATGTCATCATCACCGGCCACCGAAGCCAGTTCAGGCCCGGACGGATC 1545
Db 24 ArgMetValIleIleThrGlyProGluAlaGlnPheLysAlaGlnGlyArgIle 42

RESULT 14
US-09-764-864-1532
; Sequence 1532, Application US/09764864
```

```
; Patent No. US20020132753A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PTZ23
; CURRENT APPLICATION NUMBER: US/09/764,864
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1532
; LENGTH: 250
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (127)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (134)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-864-1532

Alignment Scores:
Pred. No.: 9-07e-08 Length: 250
Score: 19.00 Matches: 19
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.70% Indels: 0
DB: 9 Gaps: 0

US-09-270-437D-6 (1-3412) x US-09-764-864-1532 (1-250)
Qy 1630 GCTGCGCGGTGATTGGCAAGTGGCAAGCCCTGCAAGCTGCAAGCTTAACC 1686
Db 172 AlaGlyArgValIleIleGlyLysGlyGlyThrValAsnGluLeuGlnAsnLeuThr 190

RESULT 15
US-09-764-864-1114
; Sequence 1114, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PTZ23
; CURRENT APPLICATION NUMBER: US/09/764,864
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1114
; LENGTH: 261
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-764-864-1114

Alignment Scores:
Pred. No.: 9e-08 Length: 261
Score: 19.00 Matches: 19
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.70% Indels: 0
DB: 9 Gaps: 0

US-09-270-437D-6 (1-3412) x US-09-764-864-1114 (1-261)
Qy 1489 AGGATGTCATCATCACCGGCCACCGAAGCCAGTTCAGGCCCGGACGGATC 1545
Db 136 ArgMetValIleIleThrGlyProGluAlaGlnPheLysAlaGlnGlyArgIle 154

Search completed: July 16, 2004, 11:56:55
Job time : 174.5 secs
```







GenCore version 5.1.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 21, 2004, 06:04:06 ; Search time 12960 Seconds  
(without alignments)  
11410.990 Million cell updates/sec

Title: US-09-270-437D-6  
Perfect score: 3412  
Sequence: 1 ggcagcgaggagcgaggagga.....aaccttgaaaatgtttattt 3412

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 3470272 seqs, 21671516995 residues

Word size : 0

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : GenEmbl.\*

- 1: gb.ba.\*
- 2: gb.htg.\*
- 3: gb.in.\*
- 4: gb.om.\*
- 5: gb.ov.\*
- 6: gb.pat.\*
- 7: gb.ph.\*
- 8: gb.pl.\*
- 9: gb.pr.\*
- 10: gb.ro.\*
- 11: gb.sts.\*
- 12: gb.sy.\*
- 13: gb.un.\*
- 14: gb.vi.\*
- 15: em.ba.\*
- 16: em.fun.\*
- 17: em.hum.\*
- 18: em.in.\*
- 19: em.mu.\*
- 20: em.om.\*
- 21: em.or.\*
- 22: em.ov.\*
- 23: em.pat.\*
- 24: em.ph.\*
- 25: em.pl.\*
- 26: em.ro.\*
- 27: em.sts.\*
- 28: em.un.\*
- 29: em.vi.\*
- 30: em.htg.hum.\*
- 31: em.htg.inv.\*
- 32: em.htg.other.\*
- 33: em.htg.mus.\*
- 34: em.htg.pin.\*
- 35: em.htg.rod.\*
- 36: em.htg.mam.\*
- 37: em.htg.vrt.\*
- 38: em.sy.\*
- 39: em.htgo.hum.\*
- 40: em.htgo.mus.\*
- 41: em.htgo.other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	3412	100.0	3412	6	AR171865	AR171865 Sequence
2	3412	100.0	3412	6	AR343074	AR343074 Sequence
3	3412	100.0	3412	6	BD209925	BD209925 Isolated
4	3125	91.6	3633	9	BC021290	BC021290 Homo sapi
5	2143	62.8	3283	6	AR171867	AR171867 Sequence
6	2143	62.8	3283	6	AR343076	AR343076 Sequence
7	2143	62.8	3283	6	BD209927	BD209927 Isolated
8	1977	57.9	2010	9	AF117107	AF117107 Homo sapi
9	1810	53.0	3667	9	AF057352	AF057352 Homo sapi
10	1478	43.3	2640	9	HSM806243	FX537583 Homo sapi
c 11	1478	43.3	98945	9	AC016961	AC016961 Homo sapi
c 12	1296	38.0	208273	2	AC134942	AC134942 Homo sapi
c 13	586	17.2	50320	2	AC126373	AC126373 Homo sapi
14	567	16.6	710	6	BD221443	BD221443 Human gen
15	565	16.6	50320	2	AC126373	AC126373 Homo sapi
c 16	428	12.5	222876	2	AC068379	AC068379 Homo sapi
c 17	391	11.5	27352	2	AC126372	AC126372 Homo sapi
18	390	11.4	222876	2	AC068379	AC068379 Homo sapi
19	347	10.2	364	6	AX887402	AX887402 Sequence
20	347	10.2	364	6	BD027012	BD027012 Sequence
21	285	8.4	300	6	BD220404	BD220404 Human gen
22	264	7.7	264	6	AX209284	AX209284 Sequence
23	264	7.7	364	6	AX198757	AX198757 Sequence
24	263	7.7	263	6	AX209467	AX209467 Sequence
c 25	263	7.7	263	6	AX209697	AX209697 Sequence
26	263	7.7	378	6	AX198941	AX198941 Sequence
c 27	249	7.3	146376	9	AC009247	AC009247 Homo sapi
28	249	7.3	163157	9	AC108670	AC108670 Homo sapi
29	222	6.5	222	6	AX261025	AX261025 Sequence
30	222	6.5	222	6	AX261634	AX261634 Sequence
31	167	4.9	218	6	AX261764	AX261764 Sequence
32	156	4.6	306	6	BD059612	BD059612 Secreted
c 33	150	4.4	244	11	G43967	G43967 WIAF-3591-S
34	140	4.1	243	6	AX260963	AX260963 Sequence
35	135	4.0	142971	9	AC020629	AC020629 Homo sapi
36	127	3.7	3557	10	BC023758	BC023758 Mus muscu
37	127	3.7	4953	10	BC054552	BC054552 Mus muscu
38	100	2.9	91084	9	AL596177	AL596177 Human DNA
39	100	2.9	182895	2	AC015706	AC015706 Homo sapi
c 40	93	2.7	253	6	AX261078	AX261078 Sequence
c 41	92	2.7	67375	2	AC104767	AC104767 Homo sapi
c 42	92	2.7	112259	2	AC026579	AC026579 Homo sapi
c 43	92	2.7	113201	9	AP004230	AP004230 Homo sapi
c 44	92	2.7	187226	9	AC104980	AC104980 Homo sapi
c 45	83	2.4	380	6	AX335143	AX335143 Sequence

ALIGNMENTS

RESULT 1	AR171865	Sequence	6	3412 bp	DNA	linear	PAT 17-DEC-2001
LOCUS	AR171865	Sequence	6	from patent US 6297364.			
DEFINITION	AR171865	Sequence	6	from patent US 6297364.			
ACCESSION	AR171865	Sequence	6	from patent US 6297364.			
VERSION	AR171865.1	GI:17910815					
KEYWORDS	Unknown.						
SOURCE	Unknown.						
ORGANISM	Unknown.						
REFERENCE	1 (bases 1 to 3412)						
AUTHORS	Chen,Y.-T., Gure,A., Tsang,S., Stockert,E., Jager,E., Alexander,K.						
TITLE	Isolated nucleic acid molecule encoding cancer associated antigen, the antigen itself, and uses thereof						

JOURNAL Patent: US 6297364-A 6 02-OCT-2001;		Location/Qualifiers	
FEATURES		1. .3412	
source		/organism="unknown"	
ORIGIN		/mol_type="unassigned DNA"	
Query Match		100.0%; Score 3412; DB 6; Length 3412;	
Best Local Similarity		100.0%; Pred. No. 0;	
Matches 3412; Conservative		0; Mismatches 0; Indels 0; Gaps 0;	
QY	1	GGCAGCGGAGGAGCGCGGGTACCGGGCCGGGGAGCCCGGGCTCTCGGGG	60
Db	1	GGCAGCGGAGGAGCGCGGGTACCGGGCCGGGGAGCCCGGGCTCTCGGGG	60
QY	61	AAGAGCGGATGATGAACAAAGCTTTATCATCGGGAACCTGAGCCCGCGCTCACCGCGAC	120
Db	61	AAGAGCGGATGATGAACAAAGCTTTATCATCGGGAACCTGAGCCCGCGCTCACCGCGAC	120
QY	121	GACCTCGGCGAGCTTTTGGGGACAGGAAGCTGCCCTTGGCGGGACAGGTCTCTGAAG	180
Db	121	GACCTCGGCGAGCTTTTGGGGACAGGAAGCTGCCCTTGGCGGGACAGGTCTCTGAAG	180
QY	181	TCGCGCTACGCTTCGTGGAATAACCCGACAGAACTGGGCGATCCGCGCCATCGAGACC	240
Db	181	TCGCGCTACGCTTCGTGGAATAACCCGACAGAACTGGGCGATCCGCGCCATCGAGACC	240
QY	241	CTCTCGGGTAAAGTGGAAATTCATGGGAAATCATGGAAGTTGATTCTCAGTCTCTAAA	300
Db	241	CTCTCGGGTAAAGTGGAAATTCATGGGAAATCATGGAAGTTGATTCTCAGTCTCTAAA	300
QY	301	AAGCTAAGGAGCAGGAAAATTCAGATTGGAACATCCCTCTCAGCTCAGTGGGAGGTG	360
Db	301	AAGCTAAGGAGCAGGAAAATTCAGATTGGAACATCCCTCTCAGCTCAGTGGGAGGTG	360
QY	361	TGGAATCGACTTTTGGCTCAATATGGGACAGTGGAGATGTGGAACAGTCAACACAGAC	420
Db	361	TGGAATCGACTTTTGGCTCAATATGGGACAGTGGAGATGTGGAACAGTCAACACAGAC	420
QY	421	ACAGAAACCGCGTTGTCAAGCTCATATGCAACAGAGAAAGAGCAAAATAGCCATG	480
Db	421	ACAGAAACCGCGTTGTCAAGCTCATATGCAACAGAGAAAGAGCAAAATAGCCATG	480
QY	481	GAGAAGCTTAAGCGGGCATCAGTTTGAGAACTACTCTTTCAAGATTTCCTACATCCGGAT	540
Db	481	GAGAAGCTTAAGCGGGCATCAGTTTGAGAACTACTCTTTCAAGATTTCCTACATCCGGAT	540
QY	541	GAAGAGTGAAGTCCCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCT	600
Db	541	GAAGAGTGAAGTCCCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCT	600
QY	601	GAGCAAGGCGAGCCCTTGGGGCACTTCTCAGGCGAGACAGATTGATTTCCGCTGGG	660
Db	601	GAGCAAGGCGAGCCCTTGGGGCACTTCTCAGGCGAGACAGATTGATTTCCGCTGGG	660
QY	661	ATCCTGCTCCGACCCAGTTTGTGTGCTCATCATCGAAGAGAGGCTTGACCAATAAG	720
Db	661	ATCCTGCTCCGACCCAGTTTGTGTGCTCATCATCGAAGAGAGGCTTGACCAATAAG	720
QY	721	AACATCACTAAGCAGACCCAGTCCCGGTAGATATCCATAGAAAAGAGAACTCTGGAGCT	780
Db	721	AACATCACTAAGCAGACCCAGTCCCGGTAGATATCCATAGAAAAGAGAACTCTGGAGCT	780
QY	781	GCAGAGAGCTGTACCATTCATGCCACCCAGAGGGGACTTCTTGAAGCATGCCGCATG	840
Db	781	GCAGAGAGCTGTACCATTCATGCCACCCAGAGGGGACTTCTTGAAGCATGCCGCATG	840
QY	841	ATTCTTGAATCATGCAGAAAGAGGCGAGATGAGCAAACTAGCCGAAGAGATTCTCTTG	900
Db	841	ATTCTTGAATCATGCAGAAAGAGGCGAGATGAGCAAACTAGCCGAAGAGATTCTCTTG	900
QY	901	AAAACTTGGCACAAATGGCTTGGTTGGAAGCTGATTTGGAAGAGAGGCGAGAAATTG	960

Db	901	AAAACTTGGCACAAATGGCTTGGTTGGAAGCTGATTTGGAAGAGAGGCGAGAAATTG	960
Qy	961	AAGAAATTTGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAAT	1020
Db	961	AAGAAATTTGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAAT	1020
Qy	1021	ATATACAAACCCGGAAAGAACCATCATCTGTGAAGGGGCAAGTTGAGGCTGTGCCAGTGT	1080
Db	1021	ATATACAAACCCGGAAAGAACCATCATCTGTGAAGGGGCAAGTTGAGGCTGTGCCAGTGT	1080
Qy	1081	GAGATAGAGATTTATGAAGAAAGCTGCTGAGGCTTTGAAAATGATATGCTGGCTGTAAAC	1140
Db	1081	GAGATAGAGATTTATGAAGAAAGCTGCTGAGGCTTTGAAAATGATATGCTGGCTGTAAAC	1140
Qy	1141	CAACAAGCCCAATCTGATCCAGGGTTCGAACTCAGCGCACTTGGCACTTTTCAACAGGA	1200
Db	1141	CAACAAGCCCAATCTGATCCAGGGTTCGAACTCAGCGCACTTGGCACTTTTCAACAGGA	1200
Qy	1201	CTGTCCGTGCTATCTCCACAGCAGGCGCCGCGGAGCTCCCGCTGCCCTTACCCAC	1260
Db	1201	CTGTCCGTGCTATCTCCACAGCAGGCGCCGCGGAGCTCCCGCTGCCCTTACCCAC	1260
Qy	1261	CCCTTCATACCTCGGATCTTCTCCAGCGCTGTACCCCGCTACCCCGCTTGGCCCG	1320
Db	1261	CCCTTCATACCTCGGATCTTCTCCAGCGCTGTACCCCGCTACCCCGCTTGGCCCG	1320
Qy	1321	TTCCCGCATCATCTCTTATCCAGAGCAGGAGATTTGAAATCTCTTTCATCCCAACCCAG	1380
Db	1321	TTCCCGCATCATCTCTTATCCAGAGCAGGAGATTTGAAATCTCTTTCATCCCAACCCAG	1380
Qy	1381	GCTGTGGCGCCATCATCGGGAAGAGGGGCAACATCAACAGCTGGCGAGATTCGCC	1440
Db	1381	GCTGTGGCGCCATCATCGGGAAGAGGGGCAACATCAACAGCTGGCGAGATTCGCC	1440
Qy	1441	GGAGCTCTATCAAGATTGCCCCCTGCGAAGCGCCAGACGTCAGGAAAGATGTCTATC	1500
Db	1441	GGAGCTCTATCAAGATTGCCCCCTGCGAAGCGCCAGACGTCAGGAAAGATGTCTATC	1500
Qy	1501	ATCACCGGCGCCACCGAAGCCAGTTCAAGGCGCCAGGACCGGATCTTTGGGAAACTGAAA	1560
Db	1501	ATCACCGGCGCCACCGAAGCCAGTTCAAGGCGCCAGGACCGGATCTTTGGGAAACTGAAA	1560
Qy	1561	GAGGAAATCTTCTTTTAAACCCCAAGAAAGTGAAGCTGGAAGCGCATATCAGAGTGCC	1620
Db	1561	GAGGAAATCTTCTTTTAAACCCCAAGAAAGTGAAGCTGGAAGCGCATATCAGAGTGCC	1620
Qy	1621	TCCTTCCAGCTGCGCGGCTGATTTGGCAAGAGTGCAAGCGTGAACCTGAGAGTCAAGAC	1680
Db	1621	TCCTTCCAGCTGCGCGGCTGATTTGGCAAGAGTGCAAGCGTGAACCTGAGAGTCAAGAC	1680
Qy	1681	TTAACCCAGTGCAGAAAGTCACTGCTGCTGACCAAAAGCCAGATGAAATGAGGAAGTG	1740
Db	1681	TTAACCCAGTGCAGAAAGTCACTGCTGCTGACCAAAAGCCAGATGAAATGAGGAAGTG	1740
Qy	1741	ATCGTCAGAAATTTATCGGGCACTTCTTTGCTAGCAGACTGCAAGCGCAAGATCAGGAA	1800
Db	1741	ATCGTCAGAAATTTATCGGGCACTTCTTTGCTAGCAGACTGCAAGCGCAAGATCAGGAA	1800
Qy	1801	ATTGTACAAAGCTGAAGCAGGAGCAAGAAATACCTTCAGGGAGTCCCTCAGAGCGC	1860
Db	1801	ATTGTACAAAGCTGAAGCAGGAGCAAGAAATACCTTCAGGGAGTCCCTCAGAGCGC	1860
Qy	1861	AGCAAGTGAAGGCTCCCAAGGCAAGGAGCAAGAAATCAAGGATGATGAGGCTTCCAAAC	1920
Db	1861	AGCAAGTGAAGGCTCCCAAGGCAAGGAGCAAGAAATCAAGGATGATGAGGCTTCCAAAC	1920
Qy	1921	CTGACAGAAATGAGCAAAACCGAGCCAGATCGGAGCAAAACCAAGAGACCTCTGAG	1980
Db	1921	CTGACAGAAATGAGCAAAACCGAGCCAGATCGGAGCAAAACCAAGAGACCTCTGAG	1980
Qy	1981	GAAATGAGAGTCTGCGGAGCGCGGAGGACTCTGCGGAGGCGCTTGGAAACCCAGGGGC	2040
Db	1981	GAAATGAGAGTCTGCGGAGCGCGGAGGACTCTGCGGAGGCGCTTGGAAACCCAGGGGC	2040

```
QY 2041 CGAGGAGGGCGGGGAGGTGACGAGTGTGGCCGAGAACACACGAGGCCCGGCTCCCGCC 2100
Db 2041 CGAGGAGGGCGGGGAGGTCAGCCAGGTTTGCAGAACACACGAGGCCCGGCTCCCGCC 2100
QY 2101 CCCAGGGGCTCTGCGAGGCTTCAGCCATCCACTTCACCATCCACTCGGATCTCTCTGAA 2160
Db 2101 CCCAGGGGCTCTGCGAGGCTTCAGCCATCCACTTCACCATCCACTCGGATCTCTCTGAA 2160
QY 2161 CTCCACGACGCTATCCCTTTAGTTGAACATACTAGTGAACGCTGTTCAAGGCAAGC 2220
Db 2161 CTCCACGACGCTATCCCTTTAGTTGAACATACTAGTGAACGCTGTTCAAGGCAAGC 2220
QY 2221 AAAATGCACACCTTTTCTGTGGCAATCGTCTGTACATGTGTGACATATTAGAAA 2280
Db 2221 AAAATGCACACCTTTTCTGTGGCAATCGTCTGTACATGTGTGACATATTAGAAA 2280
QY 2281 GGAAGATGTTAAGATATGTGGCTGTGGTTTACACAGGGTCCCTGCAGCGTAAATAT 2340
Db 2281 GGAAGATGTTAAGATATGTGGCTGTGGTTTACACAGGGTCCCTGCAGCGTAAATAT 2340
QY 2341 TTAGAATAATATATCAATAAATCACTAACTCCAATTTTAAATCAATTAATTTT 2400
Db 2341 TTAGAATAATATATCAATAAATCACTAACTCCAATTTTAAATCAATTAATTTT 2400
QY 2401 TTTTCTTTTAAAGAGAGCAGGCTTTCTAGACTTTAAAGATAAAGTCTTTGGGAG 2460
Db 2401 TTTTCTTTTAAAGAGAGCAGGCTTTCTAGACTTTAAAGATAAAGTCTTTGGGAG 2460
QY 2461 GTCTACGGTGTAGAGAGGAGCTTTGAGGCCACCGCACAAAATTCACCCAGAGGAAT 2520
Db 2461 GTCTACGGTGTAGAGAGGAGCTTTGAGGCCACCGCACAAAATTCACCCAGAGGAAT 2520
QY 2521 CTCGTGGAAGGACACTCACCGCAGTTCTGATCACTCTGTGTATGTCACAGAGGGATA 2580
Db 2521 CTCGTGGAAGGACACTCACCGCAGTTCTGATCACTCTGTGTATGTCACAGAGGGATA 2580
QY 2581 CGGTCTCTTGAAGAGGAACCTCTGACTCTCATGCTGCTGTCTAGCTATACACCCAT 2640
Db 2581 CGGTCTCTTGAAGAGGAACCTCTGACTCTCATGCTGCTGTCTAGCTATACACCCAT 2640
QY 2641 TCTCTTTGCTTCAACAGTTTAACTGGTTTGTGATCTGATATATTAATTTCTGTCT 2700
Db 2641 TCTCTTTGCTTCAACAGTTTAACTGGTTTGTGATCTGATATATTAATTTCTGTCT 2700
QY 2701 CTCCTGTTTATCTCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 2760
Db 2701 CTCCTGTTTATCTCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 2760
QY 2761 TTTCTCATCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 2820
Db 2761 TTTCTCATCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 2820
QY 2821 GTGCTCTGAGTATCATCACAAAAGGAAACAAAAGCGGAAACACACAAACAGCTCAA 2880
Db 2821 GTGCTCTGAGTATCATCACAAAAGGAAACAAAAGCGGAAACACACAAACAGCTCAA 2880
QY 2881 CTTACCTTGGTTTCTAAAGAAACAAGAGTCAATGGTACTTGTCTAGCTTTTGGAG 2940
Db 2881 CTTACCTTGGTTTCTAAAGAAACAAGAGTCAATGGTACTTGTCTAGCTTTTGGAG 2940
QY 2941 AGGAAACAGGAACCCACCAACCCCAATCAACCAACCAAGAAAATTTCCCAATG 3000
Db 2941 AGGAAACAGGAACCCACCAACCCCAATCAACCAACCAAGAAAATTTCCCAATG 3000
QY 3001 AAAGATGTAATTTGTCTTTTGGATTTGGTGTATAGCCATCAATATTCAGCAAAATG 3060
Db 3001 AAAGATGTAATTTGTCTTTTGGATTTGGTGTATAGCCATCAATATTCAGCAAAATG 3060
QY 3061 ATTCTCTTTTAAAAAAAATGTGGAGGAAAGTAGAAATTTACCAAGTTGTGGCC 3120
Db 3061 ATTCTCTTTTAAAAAAAATGTGGAGGAAAGTAGAAATTTACCAAGTTGTGGCC 3120
```

```
QY 3121 CAGGGCGTTAAATTCACAGATTTTTTTTAAACGAGAAAAACACACAGAGAAGCTACTCAG 3180
Db 3121 CAGGGCGTTAAATTCACAGATTTTTTTTAAACGAGAAAAACACACAGAGAAGCTACTCAG 3180
QY 3181 GTGTTTTTACCTCAGACCTTGTCTGTGTTCCTTAGAGATTTTGTAAAGCTGATAG 3240
Db 3181 GTGTTTTTACCTCAGACCTTGTCTGTGTTCCTTAGAGATTTTGTAAAGCTGATAG 3240
QY 3241 TTGGAGCATTTTTTTTTTAAATAAAATAGTTGGAAAAAAATAAGATATCAACT 3300
Db 3241 TTGGAGCATTTTTTTTTTAAATAAAATAGTTGGAAAAAAATAAGATATCAACT 3300
QY 3301 GCCAGCTTGGAGAGAGTGAAGTCCAGTCCAGTGTGCAACAGCTGTTCTGAATGTCTCCGCT 3360
Db 3301 GCCAGCTTGGAGAGAGTGAAGTCCAGTCCAGTGTGCAACAGCTGTTCTGAATGTCTCCGCT 3360
QY 3361 AGCCAAAGAACCNATATGCTCTCTTTTGGACAAACCTTGAAAAATGTTTATTT 3412
Db 3361 AGCCAAAGAACCNATATGCTCTCTTTTGGACAAACCTTGAAAAATGTTTATTT 3412

RESULT 2
AR343074
LOCUS AR343074 3412 bp DNA linear PAT 17-AUG-2003
DEFINITION Sequence 6 from patent US 6576756.
ACCESSION AR343074
VERSION AR343074.1 GI:33738476
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 3412)
AUTHORS Chen,Y.-T., Gure,A., Teang,S., Stockert,E., Jager,E., Alexander,K.
and Old,L.J.
TITLE Isolated nucleic acid molecule encoding cancer associated antigen,
the antigen itself, and uses thereof
Patent: US 6576756-A 6 10-JUN-2003;
JOURNAL Location/Qualifiers
FEATURES
source
1..3412
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Query Match 100.0%; Score 3412; DB 6; Length 3412;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3412; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGCACGGAGGAGGAGGAGCGCGGTACCGGCGCGGGGAGCGCGGGCTCTCGGGG 60
Db 1 GGCACGGAGGAGGAGGAGCGCGGTACCGGCGCGGGGAGCGCGGGCTCTCGGGG 60
QY 61 AAGAGACGGATGATGAACAAGCTTTACATCGGGAACCTGAGCCCGCGCTCACCGCCGAC 120
Db 61 AAGAGACGGATGATGAACAAGCTTTACATCGGGAACCTGAGCCCGCGCTCACCGCCGAC 120
QY 121 GACCTCCGGCAGCTTTTGGGGAAGGAGTGCCTTGGCGGGACAGCTCTGCTGAAG 180
Db 121 GACCTCCGGCAGCTTTTGGGGAAGGAGTGCCTTGGCGGGACAGCTCTGCTGAAG 180
QY 181 TCCGCTACGCTTCGTGAGCTTACCCGACAGAACTCGGCCATCCGCGCATCCAGACC 240
Db 181 TCCGCTACGCTTCGTGAGCTTACCCGACAGAACTCGGCCATCCGCGCATCCAGACC 240
QY 241 CTCGGGTAAAGTGAATTCATGGGAAATTCATGGAATTTGATTCTAGTCTCTAAA 300
Db 241 CTCGGGTAAAGTGAATTCATGGGAAATTCATGGAATTTGATTCTAGTCTCTAAA 300
QY 301 AAGCTAAGCAGAGGAAATTCAGATTCGAAACATCCCTCTCTACCTGAGTGGAGGTG 360
Db 301 AAGCTAAGCAGAGGAAATTCAGATTCGAAACATCCCTCTCTACCTGAGTGGAGGTG 360
QY 361 TTGGATGACCTTTTGGCTCAATATGGGACAGTGGAGAAATGGAACAACTCAACACAG 420
Db 361 TTGGATGACCTTTTGGCTCAATATGGGACAGTGGAGAAATGGAACAACTCAACACAG 420
```



QY 2581 CCGTCTCTTGAAGAGGAACCTGTGCTCACTCTCTGCTGTCTAGCTCATACACCCATT 2640  
Db CCGTCTCTTGAAGAGGAACCTGTGCTCACTCTCTGCTGTCTAGCTCATACACCCATT 2640  
QY 2641 TCTCTTTGCTTACACAGGTTTAAATCGGTTTGTGATATGCTATATATTTCTCTGCT 2700  
Db TCTCTTTGCTTACACAGGTTTAAATCGGTTTGTGATATGCTATATATTTCTCTGCT 2700  
QY 2701 CTCTCTGTTTACTCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 2760  
Db CTCTCTGTTTACTCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 2760  
QY 2761 TTTCTCTATCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 2820  
Db TTTCTCTATCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 2820  
QY 2821 GTGCTCTGATATCATATCACAACCAAGGAAACAAAGGAAACAAACCAAGGAAACCA 2880  
Db GTGCTCTGATATCATATCACAACCAAGGAAACAAAGGAAACAAACCAAGGAAACCA 2880  
QY 2881 CTTACACTTGGTTACTCAAAAGAAACAGAGTCAATGGTACTTGTCTAGCGTTTGGAG 2940  
Db CTTACACTTGGTTACTCAAAAGAAACAGAGTCAATGGTACTTGTCTAGCGTTTGGAG 2940  
QY 2941 AGGAAACAGGAACCCACCAACCAACCAATCAACCAACCAACCAACCAACCAACCA 3000  
Db AGGAAACAGGAACCCACCAACCAACCAATCAACCAACCAACCAACCAACCAACCA 3000  
QY 3001 AAAGAATGATTTTGTCTTTTGTGATATTTGTGATATTAAGCCATCAATATTACGCA 3060  
Db AAAGAATGATTTTGTCTTTTGTGATATTTGTGATATTAAGCCATCAATATTACGCA 3060  
QY 3061 ATTCTCTTTTAAACCAACCAATGAGGAAAGTGAATTTACCAAGTTTGTGGCC 3120  
Db ATTCTCTTTTAAACCAACCAATGAGGAAAGTGAATTTACCAAGTTTGTGGCC 3120  
QY 3121 CAGGCGTTTAAATTCACAGATTTTAAACGAGAAACACACAGAAAGTCACTCAG 3180  
Db CAGGCGTTTAAATTCACAGATTTTAAACGAGAAACACACAGAAAGTCACTCAG 3180  
QY 3181 GTGTTTAACTCAGACCTTGTCTTGTGTTTCCCTTAGAGATTTTGTAAAGCTGATAG 3240  
Db GTGTTTAACTCAGACCTTGTCTTGTGTTTCCCTTAGAGATTTTGTAAAGCTGATAG 3240  
QY 3241 TTGAGACATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 3300  
Db TTGAGACATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 3300  
QY 3301 GCCAGCTGGAGAGGTGACAGTCCAGTGTGCAACAGCTGTCTGAAATTTGTTCCGCT 3360  
Db GCCAGCTGGAGAGGTGACAGTCCAGTGTGCAACAGCTGTCTGAAATTTGTTCCGCT 3360  
QY 3361 AGCCAAGAACCAATATGCTCTTTTGGACAAACCTTTGAAATTTTATTT 3412  
Db AGCCAAGAACCAATATGCTCTTTTGGACAAACCTTTGAAATTTTATTT 3412

RESULT 3  
BD209925  
LOCUS  
DEFINITION  
Isolated nucleic acid molecules encoding cancer-associated antigens, these antigens and method of using the same.  
ACCESSION  
BD209925  
VERSION  
BD209925.1  
KEYWORDS  
JP 2002512049-A/4.  
SOURCE  
Homo sapiens  
ORGANISM  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE  
1 (bases 1 to 3412)  
AUTHORS  
Chen, Y.T., Gure, A., Tsang, S., Stockert, E., Jager, E., Knuth, A. and Old, L.J.  
TITLE  
Isolated nucleic acid molecules encoding cancer-associated

antigens, these antigens and method of using the same  
Patent: JP 2002512049-A 4 23-APR-2002;  
LUDWIG INSTITUTE FOR CANCER RESEARCH  
OS Homo sapiens (human)  
PN JP 2002512049-A/4  
PD 23-APR-2002  
PF 16-MAR-1999 JP 2000545030  
PI 17-APR-1998 US 09/061709  
PI YAO TSUNG CHEN, ALI GURE, SOLAM TSANG, ELISABETH STOCKERT, ELKE JAGER.  
PI ALEXANDER KNUTH, LLOYD J OLD  
PC C12N15/09, A61K35/12, A61K39/00, A61K39/39, A61P35/00, C07K16/32, C12N1/15,  
PC C12N1/19, C12N1/21, C12N5/10, C12P21/08, C12Q1/68, G01N33/53, G01N33/574,  
PC C12N15/00, C12N5/00  
CC n-not determined(3372)  
FH Key Location/Qualifiers  
FT source 1. .3412 /organism='Homo sapiens (human)'.  
FEATURES  
source  
1. 3412  
Location/Qualifiers  
/organism='Homo sapiens'  
/mol\_type='genomic DNA'  
/db\_xref='taxon:9606'  
ORIGIN  
Query Match 100.0%; Score 3412; DB 6; Length 3412;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 3412; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GGCAGCGAGGAGGCGAGGAGCGCGGTACCGGCGCGGGGAGCGCGGGCTCTCGGG 60  
Db 1 GGCAGCGAGGAGGCGAGGAGCGCGGTACCGGCGCGGGGAGCGCGGGCTCTCGGG 60  
QY 61 AAGAGACGATGATGAAACAGCTTTATCGGGAACTGTAGCCCGCGCTCACGCCGAC 120  
Db 61 AAGAGACGATGATGAAACAGCTTTATCATCGGAACTGTAGCCCGCGCTCACGCCGAC 120  
QY 121 GACCTCCGGAGCTTTTGGGACAGGAGTGCCTTGGCGGAGCAGGTCTGCTGAAG 180  
Db 121 GACCTCCGGAGCTTTTGGGACAGGAGTGCCTTGGCGGAGCAGGTCTGCTGAAG 180  
QY 181 TCCGCTACGCTTCTGTGGACTACCCGACAGAACTCGGCGCATCCGCGCATCGAGACC 240  
Db 181 TCCGCTACGCTTCTGTGGACTACCCGACAGAACTCGGCGCATCCGCGCATCGAGACC 240  
QY 241 CTCTCGGTAAGTGGAAATTCATGGGAAATTCATGGAAATTCATGATCTCTAAA 300  
Db 241 CTCTCGGTAAGTGGAAATTCATGGGAAATTCATGGAAATTCATGATCTCTAAA 300  
QY 301 AAGCTAAGGACAGGAAATTCAGATTCGAAACATCCCTCTCCTCAGTGGGAGTG 360  
Db 301 AAGCTAAGGACAGGAAATTCAGATTCGAAACATCCCTCTCCTCAGTGGGAGTG 360  
QY 361 TTGATGGACTTTTGGCTCAATATGGGACAGTGGGAAATTCGAAACAGTCAACACAGAC 420  
Db 361 TTGATGGACTTTTGGCTCAATATGGGACAGTGGGAAATTCGAAACAGTCAACACAGAC 420  
QY 421 ACAGAAACCGGTTTCAACGTCACATATGCAACAGAGAAAGCAAAATAGCCATG 480  
Db 421 ACAGAAACCGGTTTCAACGTCACATATGCAACAGAGAAAGCAAAATAGCCATG 480  
QY 481 GAGAAGCTAAGCGGCGCATGATTTGAGAACTTACTCTTCAAGATTTCTCATATCCCGAT 540  
Db 481 GAGAAGCTAAGCGGCGCATGATTTGAGAACTTACTCTTCAAGATTTCTCATATCCCGAT 540  
QY 541 GAAGAGGTGAGTCCCTTCCGCGGCTCAGAGCCAGCGTGGGAGCCACTCTTCCCG 600  
Db 541 GAAGAGGTGAGTCCCTTCCGCGGCTCAGAGCCAGCGTGGGAGCCACTCTTCCCG 600  
QY 601 GAGCAAGGCGACGCCCTTGGGGGCACTTCTCAGGCCAGACAGATTTATTTCCGCTCGG 660

Db 601 GAGCAAGGCGACCCCTCGGGGCACTTCTCAGGCGACAGATGATTTCCCGCTGCGG 660  
QY 661 ATCTGCTGCCACCCAGCTTTGTTGTGTCATCATCGGAAAGGAGGCTTGACCAATAAG 720  
Db 661 ATCTGCTGCCACCCAGCTTTGTTGTGTCATCATCGGAAAGGAGGCTTGACCAATAAG 720  
QY 721 AACATCACTAAGCAGACCCAGTCCCGGTTAGATATCCATAGAAAAGAGAACTCTGAGCT 780  
Db 721 AACATCACTAAGCAGACCCAGTCCCGGTTAGATATCCATAGAAAAGAGAACTCTGAGCT 780  
QY 781 GCAGAGAAGCTGTCACTCCATCCATGCCACCCAGAGGGACTTCTCAAGCATGCCCATG 840  
Db 781 GCAGAGAAGCTGTCACTCCATCCATGCCACCCAGAGGGACTTCTCAAGCATGCCCATG 840  
QY 841 ATTCTTTGAATCATGAGAAAGAGGAGATGAGACCAAACTAGCTTTGAGGATTTGAGC 900  
Db 841 ATTCTTTGAATCATGAGAAAGAGGAGATGAGACCAAACTAGCTTTGAGGATTTGAGC 900  
QY 901 AAAATCTTGGCACACAAATGCTTGGTGGAGACTGATTTGGAAGAGGAGCAAAATTTG 960  
Db 901 AAAATCTTGGCACACAAATGCTTGGTGGAGACTGATTTGGAAGAGGAGCAAAATTTG 960  
QY 961 AAGAAAATTTGAAATCAAGGAGCAAGATTAACAACTCTCATCTTTGAGGATTTGAGC 1020  
Db 961 AAGAAAATTTGAAATCAAGGAGCAAGATTAACAACTCTCATCTTTGAGGATTTGAGC 1020  
QY 1021 ATATACAAACCCGAAAGAACATCACTGTGAAGGACAGTTGAGGCTGTGCGAGTGT 1080  
Db 1021 ATATACAAACCCGAAAGAACATCACTGTGAAGGACAGTTGAGGCTGTGCGAGTGT 1080  
QY 1081 GAGATAGATTAAGAGAGCTGCGTGAGGCTTTGAAAATGATATGCTGGCTGTTAAC 1140  
Db 1081 GAGATAGATTAAGAGAGCTGCGTGAGGCTTTGAAAATGATATGCTGGCTGTTAAC 1140  
QY 1141 CAACAAAGCCAACTGATCCAGGGTTGAACCTCAGCGCACTTGGCATCTTTTCAACAGGA 1200  
Db 1141 CAACAAAGCCAACTGATCCAGGGTTGAACCTCAGCGCACTTGGCATCTTTTCAACAGGA 1200  
QY 1201 CTGTGCGTGTATCTCACAGAGGCGCCGCGAGCTCCCGGTCGCCCTTACAC 1260  
Db 1201 CTGTGCGTGTATCTCACAGAGGCGCCGCGAGCTCCCGGTCGCCCTTACAC 1260  
QY 1261 CCCTTCACTACCACTCCGATATCTCCAGCTGTACCCCATCACCAAGTTGGCCG 1320  
Db 1261 CCCTTCACTACCACTCCGATATCTCCAGCTGTACCCCATCACCAAGTTGGCCG 1320  
QY 1321 TTCCCGCATCATCTCTTATCCAGAGCAGGAGATTTGAAATCTCTTCACTCCCAACCCAG 1380  
Db 1321 TTCCCGCATCATCTCTTATCCAGAGCAGGAGATTTGAAATCTCTTCACTCCCAACCCAG 1380  
QY 1381 GCTGTGGGCGCATCATCGGAGAGAGGGGGGACATCAACAGCTGGGAGATTCGCC 1440  
Db 1381 GCTGTGGGCGCATCATCGGAGAGAGGGGGGACATCAACAGCTGGGAGATTCGCC 1440  
QY 1441 GGAGCTCTATCAAGATTTGCCCTGCGAAGGCCAGAGCTCAGCGAAAGATTTGGTATC 1500  
Db 1441 GGAGCTCTATCAAGATTTGCCCTGCGAAGGCCAGAGCTCAGCGAAAGATTTGGTATC 1500  
QY 1501 ATCAACGGGCGCACCGAAGCCAGTTCAAGGCCAGGACGATCTTTGGGAACTGAAA 1560  
Db 1501 ATCAACGGGCGCACCGAAGCCAGTTCAAGGCCAGGACGATCTTTGGGAACTGAAA 1560  
QY 1561 GAGGAAACTCTTTTAACCCCAAGAGAGTGAAGCTGAAGCGCATATCAGAGTGC 1620  
Db 1561 GAGGAAACTCTTTTAACCCCAAGAGAGTGAAGCTGAAGCGCATATCAGAGTGC 1620  
QY 1621 TCTTCCACAGCTGCGGGGTGATTGGCAAGAGGTGGCAAGCCGTTGAACGAACTGCAAGC 1680  
Db 1621 TCTTCCACAGCTGCGGGGTGATTGGCAAGAGGTGGCAAGCCGTTGAACGAACTGCAAGC 1680  
QY 1681 TTAACAGTGCAGAGTCTGTCCTGTGACCAACCGCAGATGAAATGAGGAAGTG 1740

Db 1681 TTAACAGTGCAGAGTCTGTCCTGTGACCAACGCCAGATGAAAATGAGGAAGTG 1740  
QY 1741 ATCTGTCAGAAATATCGGGCACTTCTTGTAGCCAGACTGACAGCGCAAGATCAGGGAA 1800  
Db 1741 ATCTGTCAGAAATATCGGGCACTTCTTGTAGCCAGACTGACAGCGCAAGATCAGGGAA 1800  
QY 1801 ATTGTACAAACAGTGAAGCAGCAGCAGAGCAAAATACCTTCAGGAGTCCCTTCACAGCGC 1860  
Db 1801 ATTGTACAAACAGTGAAGCAGCAGCAGAGCAAAATACCTTCAGGAGTCCCTTCACAGCGC 1860  
QY 1861 AGCAAGTGAAGTCCCAACAGGACCAAGCAAAACAAACGATGAATGTAGCCCTTCCAACAC 1920  
Db 1861 AGCAAGTGAAGTCCCAACAGGACCAAGCAAAACAAACGATGAATGTAGCCCTTCCAACAC 1920  
QY 1921 CTGACAGATGAGACCAACAGCAGCAGCAGATCGGAGCAAAACCAAGACCATCTGAG 1980  
Db 1921 CTGACAGATGAGACCAACAGCAGCAGCAGATCGGAGCAAAACCAAGACCATCTGAG 1980  
QY 1981 GAATGAGAAATCTGCGGAGGCGCCAGGAGCTCTGCGGAGGCGCTGAGAACCCCGAGGGC 2040  
Db 1981 GAATGAGAAATCTGCGGAGGCGCCAGGAGCTCTGCGGAGGCGCTGAGAACCCCGAGGGC 2040  
QY 2041 CGAGGAGGCGCGGAGAGTCAAGCCAGGTTGCGCAAGAACCAAGAGCCCGCTCCCGC 2100  
Db 2041 CGAGGAGGCGCGGAGAGTCAAGCCAGGTTGCGCAAGAACCAAGAGCCCGCTCCCGC 2100  
QY 2101 CCCAGGGCTTCTGAGGCTTCAAGCCATCACTTCAACATCCACTCGGATCTCTCTGAA 2160  
Db 2101 CCCAGGGCTTCTGAGGCTTCAAGCCATCACTTCAACATCCACTCGGATCTCTCTGAA 2160  
QY 2161 CTCCCAACAGCGTATCCCTTTTAGTTGAATCAATAGGTGAACCTGTTCAAGCCAAAGC 2220  
Db 2161 CTCCCAACAGCGTATCCCTTTTAGTTGAATCAATAGGTGAACCTGTTCAAGCCAAAGC 2220  
QY 2221 AAAATGCAACCCCTTTTCTGTGGCAATCGTCTCTGTATCATGTGTGATATAGAAA 2280  
Db 2221 AAAATGCAACCCCTTTTCTGTGGCAATCGTCTCTGTATCATGTGTGATATAGAAA 2280  
QY 2281 GGAAGATGTTAAGATATGTGGCTGTGGTTACAGAGGTGCTCAGCGGTAAATAT 2340  
Db 2281 GGAAGATGTTAAGATATGTGGCTGTGGTTACAGAGGTGCTCAGCGGTAAATAT 2340  
QY 2341 TTTAGAAATAATATCAAAATCACTCAACTCAATTTTAAATTTTAAATTTT 2400  
Db 2341 TTTAGAAATAATATCAAAATCACTCAACTCAATTTTAAATTTTAAATTTT 2400  
QY 2401 TTTTCTTTTAAAGAGAAAGCGGCTTTCTAGACTTTTAAAGAAATAAGTCTTTGGAG 2460  
Db 2401 TTTTCTTTTAAAGAGAAAGCGGCTTTTCTAGACTTTTAAAGAAATAAGTCTTTGGAG 2460  
QY 2461 GTCTCAGGTGTAGAGAGAGCTTTGAGGCGCCCGCAAAAATTCACCCAGAGGAAAT 2520  
Db 2461 GTCTCAGGTGTAGAGAGAGCTTTGAGGCGCCCGCAAAAATTCACCCAGAGGAAAT 2520  
QY 2521 CTGTCGGAAGGACACTCAGCGCAGTCTGAGTCACTGTGTATGTCAACAGAGGGATA 2580  
Db 2521 CTGTCGGAAGGACACTCAGCGCAGTCTGAGTCACTGTGTATGTCAACAGAGGGATA 2580  
QY 2581 CGGTCTCTCTTGAAGAGAACTCTGTCACTCTCTCATGCTGTAGCTCATACCCAT 2640  
Db 2581 CGGTCTCTCTTGAAGAGAACTCTGTCACTCTCTCATGCTGTAGCTCATACCCAT 2640  
QY 2641 TCTCTTTGCTTCAAGGTTTAACTGGTTTTTGCATCTGTATATATTTCTGTCT 2700  
Db 2641 TCTCTTTGCTTCAAGGTTTAACTGGTTTTTGCATCTGTATATATTTCTGTCT 2700  
QY 2701 CTCTGTGTTATCTCTCCCTCCCTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2760  
Db 2701 CTCTGTGTTATCTCTCCCTCCCTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2760  
QY 2761 TTTTCT 2820  
Db 2761 TTTTCT 2820

```
QY 2821 GTGCTCTGAGTATCACATCACACAAAGGAACAAAGCGAAACACACAAACAGCCTCAA 2880
D 2821 GTGCTCTGAGTATCACATCACACAAAGGAACAAAGCGAAACACACAAACAGCCTCAA 2880
QY 2881 CTTACACCTGGTCTACTCAAAAGAACAGAGGTCAATGGTACTTGTCTTCTAGCGTTTGGAG 2940
D 2881 CTTACACCTGGTCTACTCAAAAGAACAGAGGTCAATGGTACTTGTCTTCTAGCGTTTGGAG 2940
QY 2941 AGGAAACAGGAACCCACCAACCAACCAATCAACCAACCAACCAACCAACCAACCAATG 3000
D 2941 AGGAAACAGGAACCCACCAACCAACCAATCAACCAACCAACCAACCAACCAACCAATG 3000
QY 3001 AAAGAATGTATTTTGTCTTTTGTCTTTTGTCTTTTGTCTTTTGTCTTTTGTCTTTTGTCTTT 3060
D 3001 AAAGAATGTATTTTGTCTTTTGTCTTTTGTCTTTTGTCTTTTGTCTTTTGTCTTTTGTCTTT 3060
QY 3061 ATTCCCTTTCTTTTAAAAAATAATGGAGGAAGTAGAAATTTTCAAGGTTTGTGGCC 3120
D 3061 ATTCCCTTTCTTTTAAAAAATAATGGAGGAAGTAGAAATTTTCAAGGTTTGTGGCC 3120
QY 3121 CAGGGCGTTTAAATTCACAGATTTTAAATTCACAGATTTTAAATTCACAGATTTTAAATTCACAG 3180
D 3121 CAGGGCGTTTAAATTCACAGATTTTAAATTCACAGATTTTAAATTCACAGATTTTAAATTCACAG 3180
QY 3181 GTGTTTTTACCTCAGCACCTTGTCTTGTCTTGTCTTGTCTTGTCTTGTCTTGTCTTGTCTTGTCT 3240
D 3181 GTGTTTTTACCTCAGCACCTTGTCTTGTCTTGTCTTGTCTTGTCTTGTCTTGTCTTGTCTTGTCT 3240
QY 3241 TTGGAGCATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 3300
D 3241 TTGGAGCATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 3300
QY 3301 GCCAGCGTGGAGAGTGCAGTCCAGTGCAGTCCAGTGCAGTCCAGTGCAGTCCAGTGCAGTCCAGT 3360
D 3301 GCCAGCGTGGAGAGTGCAGTCCAGTGCAGTCCAGTGCAGTCCAGTGCAGTCCAGTGCAGTCCAGT 3360
QY 3361 AGCCAAGAACCNATATGCGCTTCTTTTGGACAAACCTTGAAATGCTTTATTT 3412
D 3361 AGCCAAGAACCNATATGCGCTTCTTTTGGACAAACCTTGAAATGCTTTATTT 3412

RESULT 4
BC021290 3633 bp mRNA linear PRI 04-OCT-2003
LOCUS Homo sapiens IGF-II mRNA-binding protein 2, mRNA (cDNA clone
DEFINITION MGC:29539 IMAGE:5090334), complete cds.
ACCESSION BC021290
VERSION BC021290.2 GI:33878041
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 3633)
Straussberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altshuler, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Sapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Schetz, T.E., Brownstein, M.J., Usdin, T.B., Toshyiuki, S.,
Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, K.J., Hale, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Vallalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahy, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S.,
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smallos, D.E.,
Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
```

TITLE  
JOURNAL  
MEDLINE  
PUBMED  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL

REMARK  
COMMENT

Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences  
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)  
2338257  
12477932  
2 (bases 1 to 3633)  
Straussberg, R.  
Direct Submission  
Submitted (14-JAN-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA  
NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
On Aug 19, 2003 this sequence version replaced gi:18204200.  
Contact: MGC help desk  
Email: [cgabbs@email.nih.gov](mailto:cgabbs@email.nih.gov)  
Tissue Procurement: ATCC  
cDNA Library Preparation: Rubin Laboratory  
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC),  
Gaithersburg, Maryland;  
Web site: <http://www.nisc.nih.gov/>  
Contact: [nisc.mgc@nigri.nih.gov](mailto:nisc.mgc@nigri.nih.gov)  
Akhter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S., Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P., Hansen, N., Ho, S.-I., Karlins, E., Kwong, P., Laric, P., Legaspi, R., Maduro, Q.L., Masello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W., Tsurgeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Series: IMAGE Plate: 39 Row: 0 Column: 21  
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 27552765.

FEATURES  
source

1..3633  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="MGC:29539 IMAGE:5090334"  
/tissue\_type="Pancreas, epithelioid carcinoma"  
/clone\_lib="NIH\_MGC\_42"  
/lab\_host="DH10B-R"  
/note="Vector: pOTB7"  
1..3633  
/gene="IMP-2"  
/note="synonym: VICKZ2"  
/db\_xref="LocusID:10644"  
68..1864  
/codon\_start=1  
/product="IGF-II mRNA-binding protein 2"  
/protein\_id="AAH21290.1"  
/db\_xref="GI:18204201"  
/db\_xref="LocusID:10644"  
/translation="MNKLYGNLSPAVTADLRLQFDRKLPLAGOVLLKSGYAFVDY  
PDQNAIRATLTSKGVHLKIMEVDYVSVKLRKIQIRNI PHLOWELDGLLA  
QGTGVNVEQVNDTETAVVNTVYATREAKIAMKLSGHQFENYSFKTSIYIDEEVS  
SPSPQRAQRGDHSSREOQHAGTQSOAQIDPFLILYPTQFVGAIIQKGGITIKNI  
TKYQSRQVDIHRKENSAGAEKVTHATPEGTSEACRMILEMQKADTKKAEIPL  
KILANGLVRLIKREGRNKKIHEHTGKTISSLDLSIYNPTITVKGVEACA  
SAEIEIMKRLREAFENMLAVQANLIPGLNLSALGIFSTGLSPFAPGRGAPPA  
APYHPFTTHSGYFSLYPHPQFPFHHSYEQEIVNFIPTQAVGAILGKGAHK  
QUARFAGASIKIAPAEQPDVSERMWIITGPPEAQKAGRIIPTQAKIENFFPKBEVK  
LEAHIRVPSSTAGRVIGKGVNQLNTSEVIVPRQDTPENEEVIVRIIGHFFA  
SQTQAKRIEIVQVQKQEQKYPQGVASQSK"

gene

CDS

misc\_feature

/note="RRM; Region: RNA recognition motif"  
/db\_xref="CDD:smart00360"  
650..853



/note="KH; Region: K homology RNA-binding domain"			
/db_xref="CDD:smart00322"			
1340..1555			
/note="KH; Region: K homology RNA-binding domain"			
/db_xref="CDD:smart00322"			
ORIGIN			
Query Match 91.6%; Score 3125; DB 9; Length 3633;			
Best Local Similarity 99.9%; Pred. No. 0;			
Matches 3365; Conservative 0; Mismatches 0; Indels 2; Gaps 2;			
QY	6	CGGAGAGGCGAGGACGCGCGGTACCGGGCCGGGGAGCCGGGCTCTCGGGGAAGAG	65
Db	1	CGGAGAGGCGAGGACGCGCGGTACCGGGCCGGGGAGCCGGGCTCTCGGGGAAGAG	60
QY	66	ACGGATGATCAACAAGCTTTACATCGGGAACCTGAGCCCGCGCTCACCGCCGACACCT	125
Db	61	ACGGATGATGACAAGCTTTACATCGGGAACCTGAGCCCGCGCTCACCGCCGACACCT	120
QY	126	CCGGCAGCTCTTTGGGGACAGGAAGCTGCCCTTGGCGGACAGGTCTCTGCTGAAGTCCGG	185
Db	121	CCGGCAGCTCTTTGGGGACAGGAAGCTGCCCTTGGCGGACAGGTCTCTGCTGAAGTCCGG	180
QY	186	CTACGCTTTCTGTGACTACCCGACAGAACTGGGCCATCCGCGCCATCGAGACCTCTCTC	245
Db	181	CTACGCTTTCTGTGACTACCCGACAGAACTGGGCCATCCGCGCCATCGAGACCTCTCTC	240
QY	246	GGGTAAAGTGGAAATTCATGGAAGTGTGATCTCAGTCTCTTAAAGACT	305
Db	241	GGGTAAAGTGGAAATTCATGGAAGTGTGATCTCAGTCTCTTAAAGACT	300
QY	306	AAGGAGCAGGAAATTCAGATTCCGAACTCCCTCCTCAGTGGGAGGTGTGGA	365
Db	301	AAGGAGCAGGAAATTCAGATTCCGAACTCCCTCCTCAGTGGGAGGTGTGGA	360
QY	366	TGGACTTTTGGCTCAATATGGGACAGTGGAGAAATGTGAAACAAGTCAACACAGACACAGA	425
Db	361	TGGACTTTTGGCTCAATATGGGACAGTGGAGAAATGTGAAACAAGTCAACACAGACACAGA	420
QY	426	AACCGCGTGTTCACAGTTCACATATGCAACAAGAGAGAAATAGCCATGGAGAA	485
Db	421	AACCGCGTGTTCACAGTTCACATATGCAACAAGAGAGAAATAGCCATGGAGAA	480
QY	486	GCTAAGCGGGCATCAGTTTGAGAACTACTCTTCAAGATTTCCTACATCCCGGATGAGA	545
Db	481	GCTAAGCGGGCATCAGTTTGAGAACTACTCTTCAAGATTTCCTACATCCCGGATGAGA	540
QY	546	GGTGAAGTCCCTTTCGCCCCCTCAGCGAGCCGAGCGTGGGAGCACTCTTCCCGGAGCA	605
Db	541	GGTGAAGTCCCTTTCGCCCCCTCAGCGAGCCGAGCGTGGGAGCACTCTTCCCGGAGCA	600
QY	606	AGGCCACGCCCTTGGGGGCACTTCTCAGGCCAGACAGATTGATTTCCGCTGCGGATCT	665
Db	601	AGGCCACGCCCTTGGGGGCACTTCTCAGGCCAGACAGATTGATTTCCGCTGCGGATCT	660
QY	666	GGTCCCCCAGCTTTGTTGGTGCCATCATCGGAAAGAGGGCTTGACCATAAAGAACAT	725
Db	661	GGTCCCCCAGCTTTGTTGGTGCCATCATCGGAAAGAGGGCTTGACCATAAAGAACAT	720
QY	726	CACATAAGCAGACCCAGTCCGGGTAGATATCCATAGAAAAGAACTCTGGAGCTGCAGA	785
Db	721	CACATAAGCAGACCCAGTCCGGGTAGATATCCATAGAAAAGAACTCTGGAGCTGCAGA	780
QY	786	GAAAGCTGTACCATCCATCCACCCAGAGGGGACTTCTGAAAGCATGCGCATGATCT	845
Db	781	GAAAGCTGTACCATCCATCCATCCACCCAGAGGGGACTTCTGAAAGCATGCGCATGATCT	840
QY	846	TGAAATCATGCGAAAGAGGCGATGAGACAAACTAGCCGAGAGATTCCTCTGAAAT	905
Db	841	TGAAATCATGCGAAAGAGGCGATGAGACAAACTAGCCGAGAGATTCCTCTGAAAT	900
QY	906	CTTGGCACAATGGCTTGGTTCGAAAGCTGATTTGAAAAGAGGCGAGAAATTTGAAGAA	965

Db	901	CTTGGCACACAATGGCTTGGTTGGAAGACTGATTGGAAAAGAGCGAGAAATTTGAAGAA	960
QY	966	AATTGAACATGAACACAGGACCAAGATAAACAATCTCATCTTTGACGATTTGAGCATATA	1025
Db	961	AATTGAACATGAACACAGGACCAAGATAAACAATCTCATCTTTGACGATTTGAGCATATA	1020
QY	1026	CAACCCGGAAAGAACCATCACTGTGAAGGGCACAGTTGAGGCGCTGTGCCAGTGTGAGAT	1085
Db	1021	CAACCCGGAAAGAACCATCACTGTGAAGGGCACAGTTGAGGCGCTGTGCCAGTGTGAGAT	1080
QY	1086	AGAGATTATGAAGAGCTGCGTGGAGCCCTTTGAAAATGATATGCTGGCTTTAAACCAACA	1145
Db	1081	AGAGATTATGAAGAGCTGCGTGGAGCCCTTTGAAAATGATATGCTGGCTTTAAACCAACA	1140
QY	1146	AGCCAACTGTATCCAGAGGTTGAACTCAGCGCACTTGGCATCTTTTCAACAGGACTGTC	1205
Db	1141	AGCCAACTGTATCCAGAGGTTGAACTCAGCGCACTTGGCATCTTTTCAACAGGACTGTC	1200
QY	1206	CGTGTCTATCTCCACAGCAGGCGCCCGGGAGCTCCCGCGCTGCCCTTACACCCCTT	1265
Db	1201	CGTGTCTATCTCCACAGCAGGCGCCCGGGAGCTCCCGCGCTGCCCTTACACCCCTT	1260
QY	1266	CACTACCCACTCGGGATCTTCTCCAGCCCTGTACCCCATCACAGTTTGGCCCGTCCG	1325
Db	1261	CACTACCCACTCGGGATCTTCTCCAGCCCTGTACCCCATCACAGTTTGGCCCGTCCG	1320
QY	1326	GCATCATCACTCTTATCCAGAGCAGGAGATTGTGAATCTCTTTCATCCCAACCCAGCTGT	1385
Db	1321	GCATCATCACTCTTATCCAGAGCAGGAGATTGTGAATCTCTTTCATCCCAACCCAGCTGT	1380
QY	1386	GGCGCGCATCATCGGGAAGAGGGGCAACATCAAAACAGCTGGCGAGATTCCCGGAGC	1445
Db	1381	GGCGCGCATCATCGGGAAGAGGGGCAACATCAAAACAGCTGGCGAGATTCCCGGAGC	1440
QY	1446	CTTATCAAGATTGCCCTCGCGAAGCGCCAGACGTCAGGAAAGGATGTCATCATCAC	1505
Db	1441	CTTATCAAGATTGCCCTCGCGAAGCGCCAGACGTCAGGAAAGGATGTCATCATCAC	1500
QY	1506	CGGCGCACCGGAGCCCGAGTTCAAGGCCCAAGGACCGGATCTTTGGGAAACTGAAAGAGGA	1565
Db	1501	CGGCGCACCGGAGCCCGAGTTCAAGGCCCAAGGACCGGATCTTTGGGAAACTGAAAGAGGA	1560
QY	1566	AAACTCTTTTAAACCCCAAGAGTGAAGCTGGAAGCGCATATCAGAGTGCCTCTTC	1625
Db	1561	AAACTCTTTTAAACCCCAAGAGTGAAGCTGGAAGCGCATATCAGAGTGCCTCTTC	1620
QY	1626	CACAGCTGGCGGGTGATGCAAGGTGCAAGCGTGAACGAACTGACAGAACTTAAAC	1685
Db	1621	CACAGCTGGCGGGTGATGCAAGGTGCAAGCGTGAACGAACTGACAGAACTTAAAC	1680
QY	1686	CAGTGCAGAACTCATCGTGCCTCGTGACCAACCCAGATGAAATGAGGAAGTATCGT	1745
Db	1681	CAGTGCAGAACTCATCGTGCCTCGTGACCAACCCAGATGAAATGAGGAAGTATCGT	1740
QY	1746	CAGAATTATCGGGCACTTCTTTGCTAGCCAGCTGCACAGCGCAGATCAGGAAATGT	1805
Db	1741	CAGAATTATCGGGCACTTCTTTGCTAGCCAGCTGCACAGCGCAGATCAGGAAATGT	1800
QY	1806	ACAACAGTGAAGCAGCAGGAGCAGAAATACCTCAGGAGTGCCTCACAACCGAGCAA	1865
Db	1801	ACAACAGTGAAGCAGCAGGAGCAGAAATACCTCAGGAGTGCCTCACAACCGAGCAA	1860
QY	1866	GTGAGCTCCCAACAGCACCAGCAAAACCAACGATGAAATGAGCCCTTCCAAACCTGAC	1925
Db	1861	GTGAGCTCCCAACAGCACCAGCAAAACCAACGATGAAATGAGCCCTTCCAAACCTGAC	1920
QY	1926	AGAATGAGACCAACAGCAGCAGCAGATCGGGAGCAACCAAGAACCTCTGAGGAATG	1985
Db	1921	AGAATGAGACCAACAGCAGCAGCAGATCGGGAGCAACCAAGAACCTCTGAGGAATG	1980
QY	1986	AGAAGTCTGCGAGCGCGCAGGGACTCTGCGAGGCGCTTGAGAACCCCGAGGGCGGAGG	2045
Db	1981	AGAAGTCTGCGAGCGCGCAGGGACTCTGCGAGGCGCTTGAGAACCCCGAGGGCGGAGG	2040



QY 2046 AGGGGGGGAAGGTGAGCAGGTTTGGCAGAACCCAGAGCCCGGCTCCGCGCCCCCA 2105  
Db |||||||  
QY 2041 AGGGGGGGAAGGTGAGCAGGTTTGGCAGAACCCAGAGCCCGGCTCCGCGCCCCCA 2100  
Db |||||||  
QY 2106 GGGCTTCTGAGGCTTTCAGCCATCCACTTCACCATCCACTCGGATCTCTCTGAATCCC 2165  
Db |||||||  
QY 2101 GGGCTTCTGAGGCTTTCAGCCATCCACTTCACCATCCACTCGGATCTCTCTGAATCCC 2160  
Db |||||||  
QY 2166 ACCAGCCTATCCCTTTAGTTGAATCAATAGGTGAAGCTGTTCAAGCAAGCAAAAT 2225  
Db |||||||  
QY 2161 ACAGCCTATCCCTTTAGTTGAATCAATAGGTGAAGCTGTTCAAGCAAGCAAAAT 2220  
Db |||||||  
QY 2226 GCACACCTTTTCTGCTGGCAATCTCTCTGACATGTTGACATATTAAGAAAGGAA 2285  
Db |||||||  
QY 2221 GCACACCTTTTCTGCTGGCAATCTCTCTGACATGTTGACATATTAAGAAAGGAA 2280  
Db |||||||  
QY 2286 GATGTTAAGATATGTGGCTGTGGGTTACACAGGTCCTGTCAGCGGTAAATATTTAG 2345  
Db |||||||  
QY 2281 GATGTTAAGATATGTGGCTGTGGGTTACACAGGTCCTGTCAGCGGTAAATATTTAG 2340  
Db |||||||  
QY 2346 AAATAATATCAATAAATCACTCACTCACTCACTCACTCACTCACTCACTCACTCACT 2405  
Db |||||||  
QY 2341 AAATAATATCAATAAATCACTCACTCACTCACTCACTCACTCACTCACTCACTCACT 2400  
Db |||||||  
QY 2406 CTTTTTAAAGAGAACAGCGCTTTCTGACACTTTTAAAGATTAAGTCTTTGGAGGTCTC 2465  
Db |||||||  
QY 2401 CTTTTTAAAGAGAACAGCGCTTTCTGACACTTTTAAAGATTAAGTCTTTGGAGGTCTC 2460  
Db |||||||  
QY 2466 AGGTGTAGAGAGGCTTTGAGGCCACCCGACACAAATTCACCCAGAGGAAATCTGT 2525  
Db |||||||  
QY 2461 AGGTGTAGAGAGGCTTTGAGGCCACCCGACACAAATTCACCCAGAGGAAATCTGT 2520  
Db |||||||  
QY 2526 CGAAGGACACTCAGCGAGTTCTGGATCACTGATGTATGTCAACAGAGGATACCGTC 2585  
Db |||||||  
QY 2521 CGAAGGACACTCAGCGAGTTCTGGATCACTGATGTATGTCAACAGAGGATACCGTC 2580  
Db |||||||  
QY 2586 TCCTTGAAGAGAAACTCTGTCACTCCTCATCCCTGCTAGCTCATACACCAATTTCT 2645  
Db |||||||  
QY 2581 TCCTTGAAGAGAAACTCTGTCACTCCTCATCCCTGCTAGCTCATACACCAATTTCT 2640  
Db |||||||  
QY 2646 TTGCTTCAAGGTTTAACTGGTGTGTCATCTGCTATATATTTCTGTCTCTCTC 2705  
Db |||||||  
QY 2641 TTGCTTCAAGGTTTAACTGGTGTGTCATCTGCTATATATTTCTGTCTCTCTC 2700  
Db |||||||  
QY 2706 TGTATATCTCTCCCTCCCTCCCTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2765  
Db |||||||  
QY 2701 TGTATATCTCTCCCTCCCTCCCTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2760  
Db |||||||  
QY 2766 TCATCCCTCCATCTCAATCCCGTATCTACGACCCCGCCCGCCCGCCAGGCAAGGAGTCT 2825  
Db |||||||  
QY 2761 TCATCCCTCCATCTCAATCCCGTATCTACGCA - CCCCCCCCCCCCCAGGCAAGGAGTCT 2819  
Db |||||||  
QY 2826 CTGAGTATCATCATCAACAAAGGAAACAAAGCGAAACACACAAACAGCCCTCAACTTAC 2885  
Db |||||||  
QY 2820 CTGAGTATCATCATCAACAAAGGAAACAAAGCGAAACACACAAACAGCCCTCAACTTAC 2879  
Db |||||||  
QY 2886 ACTTGGTTACTCAAAAGAACAGAGTCAATGGTACTTGTCTAGCGTTTGGAAAGAGAA 2945  
Db |||||||  
QY 2880 ACTTGGTTACTCAAAAGAACAGAGTCAATGGTACTTGTCTAGCGTTTGGAAAGAGAA 2939  
Db |||||||  
QY 2946 AACAGGAACCCACCAACCAACCAATCAACCAACCAAGAAATTTCCCAATGAAGA 3005  
Db |||||||  
QY 2940 AACAGGAACCCACCAACCAACCAATCAACCAACCAAGAAATTTCCCAATGAAGA 2999  
Db |||||||  
QY 3006 ATGTATTTTGTCTTTTGTGTTATAGCCATCAATATTCAGCAAAATGATTC 3065  
Db |||||||  
QY 3000 ATGTATTTTGTCTTTTGTGTTATAGCCATCAATATTCAGCAAAATGATTC 3059  
Db |||||||  
QY 3066 TTTCTTT - AAAAAAATAATGTGAGGAAAGTAGAATTTACAGGTTTGTGGCCAGG 3124  
Db |||||||  
QY 3060 TTTCTTTAAAAAATAATGTGAGGAAAGTAGAATTTACCAAGGTTGTGGCCAGG 3119  
Db |||||||

QY 3125 GCCTTAAATTCACAGATTTTTTTTAAACGAGAAAAACACACAGAACAGCTACCTCAGTGT 3184  
Db |||||||  
QY 3120 GCCTTAAATTCACAGATTTTTTTTAAACGAGAAAAACACACAGAACAGCTACCTCAGTGT 3179  
Db |||||||  
QY 3185 TTTTACCTCAGCACTTGTCTTGTGTTCCCTTAGAGATTTTGTAAAGCTGATAGTTGG 3244  
Db |||||||  
QY 3180 TTTTACCTCAGCACTTGTCTTGTGTTCCCTTAGAGATTTTGTAAAGCTGATAGTTGG 3239  
Db |||||||  
QY 3245 AGCATTTTTTTTATTTTTTAAATAAATAAGTTGGAAAAAATAAGATATCAACTGCCA 3304  
Db |||||||  
QY 3240 AGCATTTTTTTTATTTTTTAAATAAATAAGTTGGAAAAAATAAGATATCAACTGCCA 3299  
Db |||||||  
QY 3305 GCTCGAGAGGTGACAGTCCCAAGTGTCAACAGCTGTCTGTAATTTCTTCCGTAGCC 3364  
Db |||||||  
QY 3300 GCTCGAGAGGTGACAGTCCCAAGTGTCAACAGCTGTCTGTAATTTCTTCCGTAGCC 3359  
Db |||||||  
QY 3365 AAGAAC 3371  
Db |||||||  
QY 3360 AAGAAC 3366  
Db |||||||

RESULT 5  
ARI171867  
LOCUS ARI171867 3283 bp DNA linear PAT 17-DEC-2001  
DEFINITION Sequence 8 from patent US 6297364.  
ACCESSION ARI171867  
VERSION ARI171867.1 GI:17910817  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 3283)  
AUTHORS Chen,Y.-T., Gure,A., Tsang,S., Stockert,E., Jager,E., Alexander,K.  
and Old,L.J.  
TITLE Isolated nucleic acid molecule encoding cancer associated antigen,  
the antigen itself, and uses thereof  
JOURNAL Patent: US 6297364-A 8 02-OCT-2001;  
FEATURES Location/Qualifiers  
source 1..3283  
/organism="unknown"  
/mol\_type="unassigned DNA"

Query Match 62.8%; Score 2143; DB 6; Length 3283;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1270 ACCCACTCCGATATCTTCCAGCTGTACCCCATCACCAAGTTTGGCCCTTCCCGCAT 1329  
Db |||||||  
QY 1141 ACCCACTCCGATATCTTCCAGCTGTACCCCATCACCAAGTTTGGCCCTTCCCGCAT 1200  
Db |||||||  
QY 1330 CATCACTCTTATCCAGAGCAGGAGATTGTGAATCTTTCATCCCAACCCAGGCTGTGGC 1389  
Db |||||||  
QY 1201 CATCACTCTTATCCAGAGCAGGAGATTGTGAATCTTTCATCCCAACCCAGGCTGTGGC 1260  
Db |||||||  
QY 1390 GCCATCATCGGAGAAAGGGGCAACATCAACAGCTGGCGAGATTTCGGCGAGCTCT 1449  
Db |||||||  
QY 1261 GCCATCATCGGAGAAAGGGGCAACATCAACAGCTGGCGAGATTTCGGCGAGCTCT 1320  
Db |||||||  
QY 1450 ATCAAGATTGCCCTCTCGGAAGCCACAGCTCAGCGAAAGGATGTCATCATCACCGGG 1509  
Db |||||||  
QY 1321 ATCAAGATTGCCCTCTCGGAAGCCACAGCTCAGCGAAAGGATGTCATCATCACCGGG 1380  
Db |||||||  
QY 1510 CCACCGAAGCCAGTTTCAGGCCCGAGGACGATCTTTGGGAACTGAAAGAGGAAAC 1569  
Db |||||||  
QY 1381 CCACCGAAGCCAGTTTCAGGCCCGAGGACGATCTTTGGGAACTGAAAGAGGAAAC 1440  
Db |||||||  
QY 1570 TTTCTTAAACCCCAAGAAAGTGAAGCTGGAAGCCATATCAGAGTGCCTCTTCCACA 1629  
Db |||||||  
QY 1441 TTTCTTAAACCCCAAGAAAGTGAAGCTGGAAGCCATATCAGAGTGCCTCTTCCACA 1500  
Db |||||||  
QY 1630 GCTGGCGGGTGTATGGCAAGGTGGCAAGCCGTGAACGAACTGCAGAACTTAAACAGT 1689  
Db |||||||



|||||  
1141 ACCCACTCGGAACTCTTCAGCCTGTACCCCATACAGTTTGGCCGCTTCGCGCAT 1200  
QY CATCACTTTATCCAGAGCAGGAGATTGTGAATCTCTTCATCCCAACCCAGGCTGTGGGC 1389  
Db CATCACTTTATCCAGAGCAGGAGATTGTGAATCTCTTCATCCCAACCCAGGCTGTGGGC 1260  
QY GCATCATCGGGAAGAGGGGCGACATCAAAACAGCTGGCGAGATTGCGCGGAGCCTCT 1449  
Db GCCATCATCGGGAAGAGGGGCGACATCAAAACAGCTGGCGAGATTGCGCGGAGCCTCT 1320  
QY ATCAAGATTGCCCTTGGCGAAGGCCAGACGTCAGCGAAAGGATGGTCAATCATCACCCGG 1509  
Db ATCAAGATTGCCCTTGGCGAAGGCCAGACGTCAGCGAAAGGATGGTCAATCATCACCCGG 1380  
QY CAACCGGAAGCCAGTTCAAGGCCCAGGGACGGATCTTTGGGAAACTGAAAGAGGAAAC 1569  
Db CAACCGGAAGCCAGTTCAAGGCCCAGGGACGGATCTTTGGGAAACTGAAAGAGGAAAC 1440  
QY TTCTTTAAACCCCAAGAGAAAGTGAAGCTGGGAAGCGCATATCAGAGTGCCTCTTCCACA 1629  
Db TTCTTTAAACCCCAAGAGAAAGTGAAGCTGGGAAGCGCATATCAGAGTGCCTCTTCCACA 1620  
QY GTGCGCGGGTGATTGGCAAGGTGGCAAGACCGTGAAACGACTGCAGAACTTAAACCACT 1689  
Db GTGCGCGGGTGATTGGCAAGGTGGCAAGACCGTGAAACGACTGCAGAACTTAAACCACT 1560  
QY GGAGAAGTCATGTCGCTGTGACCAAAACCGCAGATGAAATGAGGAAAGTATGTCAGA 1749  
Db GGAGAAGTCATGTCGCTGTGACCAAAACCGCAGATGAAATGAGGAAAGTATGTCAGA 1620  
QY ATTATCGGCACTTTCTTTGCTAGCCAGACTGCACAGCGCAAGATCAGGGAATTTGTACAA 1809  
Db ATTATCGGCACTTTCTTTGCTAGCCAGACTGCACAGCGCAAGATCAGGGAATTTGTACAA 1680  
QY CAGGTGAAGCAGCAGGAGCAGAAATACCTCAGGAGTGCCTTCACAGCGCAGCAAGTGA 1869  
Db CAGGTGAAGCAGCAGGAGCAGAAATACCTCAGGAGTGCCTTCACAGCGCAGCAAGTGA 1740  
QY GGCTCCCAAGCAACAGCAAAACAAACCGATGAATGTAGCCCTTCCAAACCTGCAGAA 1929  
Db GGCTCCCAAGCAACAGCAAAACAAACCGATGAATGTAGCCCTTCCAAACCTGCAGAA 1800  
QY TGAGACCAAAACGACGCCAGATCGGAGCAAAACCAAGCAATCTGAGGAATGAGAA 1989  
Db TGAGACCAAAACGACGCCAGATCGGAGCAAAACCAAGCAATCTGAGGAATGAGAA 1860  
QY GTCTCGGAGGCGGCGAGGACTCTGCCGAGGCCCTGAGAACCCAGGCGCGAGGAGG 2049  
Db GTCTCGGAGGCGGCGAGGACTCTGCCGAGGCCCTGAGAACCCAGGCGCGAGGAGG 1920  
QY GCGGGAAGGTCAGCCAGGTTTGCAGAACCAACGAGGCCGCTCCGCCCCCAGGCG 2109  
Db GCGGGAAGGTCAGCCAGGTTTGCAGAACCAACGAGGCCGCTCCGCCCCCAGGCG 1980  
QY TTCTCAGGCTTCAGCCATCCATTCACATCCATCGGATCTCTCCTGAACTCCACGA 2169  
Db TTCTCAGGCTTCAGCCATCCATTCACATCCATCGGATCTCTCCTGAACTCCACGA 2040  
QY CGCTATCCCTTTAGTTGAACCTAAACATAGTGAACGTTTCAAGCCAGCAAGCAAAATGCAC 2229  
Db CGCTATCCCTTTAGTTGAACCTAAACATAGTGAACGTTTCAAGCCAGCAAGCAAAATGCAC 2100  
QY ACCCTTTTCTGTGCAAAATCGTCTGTACATGTGTGTACATATAGAAAGGGAAGATG 2289  
Db ACCCTTTTCTGTGCAAAATCGTCTGTACATGTGTGTACATATAGAAAGGGAAGATG 2160  
QY TTAAGATATGTGGCTGTGGGTTCACAGGTCCTGACGGGTAAATATATTTTGAAT 2349  
Db TTAAGATATGTGGCTGTGGGTTCACAGGTCCTGACGGGTAAATATATTTTGAAT 2220  
QY AATATATCAAAATACTCAACTAACTCCAAATTTTAAATCAATTAATTAATTTTCTTT 2409  
|||||

Db 2221 AATATATCAAAATACTCAACTAACTCCAAATTTTAAATCAATTAATTAATTTTCTTT 2280  
QY TTAAGAGAGAAAGCAGGCTTTTCTAGACTTTAAAGAAATAAAGTCTTTGGGAGTCTCACGG 2469  
Db TTAAGAGAGAAAGCAGGCTTTTCTAGACTTTAAAGAAATAAAGTCTTTGGGAGTCTCACGG 2340  
QY TGTAGAGAGAGGCTTTTGAGGGCACCCGACAAAATTCACCCAGAGGGAATCTCGTCGGA 2529  
Db TGTAGAGAGAGGCTTTTGAGGGCACCCGACAAAATTCACCCAGAGGGAATCTCGTCGGA 2400  
QY AGGACACTCAGCGCAGTCTTGATCACTGCTGTGTATGTCAACAGAGGAGTACCGTCTCT 2589  
Db AGGACACTCAGCGCAGTCTTGATCACTGCTGTGTATGTCAACAGAGGAGTACCGTCTCT 2460  
QY TGAAGAGGAAACTCTGTCTCACTCTCTCATGCTGTCTAGCTCATACACCAATTTCTCTTTCG 2649  
Db TGAAGAGGAAACTCTGTCTCACTCTCTCATGCTGTCTAGCTCATACACCAATTTCTCTTTCG 2520  
QY TTCACAGGTTTAAACTGGTTTTTGGCATACTGCTATATAAATCTCTCTCTCTCTCT 2709  
Db TTCACAGGTTTAAACTGGTTTTTGGCATACTGCTATATAAATCTCTCTCTCTCTCTCT 2580  
QY TATCTCTCCCTCCCTCCCTCTCTCTCCATCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2769  
Db TATCTCTCCCTCCCTCCCTCTCTCTCCATCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2640  
QY CCTCCTCATCTCAATCCGTTATCTAGCGACCCCTCCCTCCCTCCCTCCCTCCCTCTCT 2829  
Db CCTCCTCATCTCAATCCGTTATCTAGCGACCCCTCCCTCCCTCCCTCCCTCCCTCTCT 2700  
QY GTATCATATCATCAAAAAGGAAACAAAAGGAAACACACAAAACAGCCTCAACTTACAT 2889  
Db GTATCATATCATCAAAAAGGAAACAAAAGGAAACACACAAAACAGCCTCAACTTACAT 2760  
QY GTTACTCAAAAGAACAAAGAGTCAATGTACTGTCTCTAGCGTTTGGAGAGGAAACA 2949  
Db GTTACTCAAAAGAACAAAGAGTCAATGTACTGTCTCTAGCGTTTGGAGAGGAAACA 2820  
QY GGAACCCCAACCAACCAATCAACCAACCAAGAAAGAAATTTCCCAATGAAAGAAATGT 3009  
Db GGAACCCCAACCAACCAATCAACCAACCAAGAAAGAAATTTCCCAATGAAAGAAATGT 2880  
QY ATTTTGTCTTTTGGTATAGGCTCAATATTAAGCCATCAATATTCAGCAAAATGATTCCTTC 3069  
Db ATTTTGTCTTTTGGTATAGGCTCAATATTAAGCCATCAATATTCAGCAAAATGATTCCTTC 2940  
QY TTTTAAAAAAAATGTGGAGAAAGTAGAAATTTTACCAAGTTCTGTGCCCCAGGCGTT 3129  
Db TTTTAAAAAAAATGTGGAGAAAGTAGAAATTTTACCAAGTTCTGTGCCCCAGGCGTT 3000  
QY AAATTCACAGATTTTTPAACGAGAAAAACACACAGAAAGAGTACCTCAGGTGTTTTTA 3189  
Db AAATTCACAGATTTTTPAACGAGAAAAACACACAGAAAGAGTACCTCAGGTGTTTTTA 3060  
QY CCTCAGCAGCTTGTCTTGTGTCTTCCCTTAGAGATTTTGTAAAGCTGTAGTTGAGCAT 3249  
Db CCTCAGCAGCTTGTCTTGTGTCTTCCCTTAGAGATTTTGTAAAGCTGTAGTTGAGCAT 3120  
QY TTTTATTTTATTTTAAATAAATGAGTTGAAAAAATAAGATATCACTGCCAGCCTG 3309  
Db TTTTATTTTATTTTAAATAAATGAGTTGAAAAAATAAGATATCACTGCCAGCCTG 3180  
QY GAGAAGGTGACAGTCCAAGTGTCAACAGCTGTTCTGAATTTGTCTTCGCTAGCAAGAA 3369  
Db GAGAAGGTGACAGTCCAAGTGTCAACAGCTGTTCTGAATTTGTCTTCGCTAGCAAGAA 3240  
QY CCNATATGGCCCTCTTTTGGCAAAACCTTGAAAAATGTTTATTT 3412  
Db CCNATATGGCCCTCTTTTGGCAAAACCTTGAAAAATGTTTATTT 3283

LOCUS BD209927 3283 bp DNA linear PAT 17-JUL-2003  
DEFINITION Isolated nucleic acid molecules encoding cancer-associated  
antigens, these antigens and method of using the same.  
ACCESSION BD209927  
VERSION BD209927.1 GI:33019697  
KEYWORDS JP 2002512049-A/6.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 3283)  
AUTHORS Chen,Y.T., Gure,A., Teang,S., Stockert,E., Jager,E., Knuth,A. and  
Old,L.J.  
TITLE Isolated nucleic acid molecules encoding cancer-associated  
antigens, these antigens and method of using the same  
JOURNAL Patent: JP 2002512049-A 6 23-APR-2002;  
COMMENT LUDWIG INSTITUTE FOR CANCER RESEARCH  
OS Homo sapiens (human)  
PN JP 2002512049-A/6  
PD 23-APR-2002  
PF 16-MAR-1999 JP 2000545030  
PI 17-APR-1998 US 09/061709  
PR YAO TSENG CHEN,ALI GURE,SOLAM TSANG,ELISABETH STOCKERT,ELKE  
PI JAGER,  
PI ALEXANDER KNUTH,LLOYD J OLD  
PC C12N15/09,A61K35/12,A61K39/00,A61K39/39,A61P35/00,C07K16/32,  
PC C12N1/15,  
PC C12N1/19,C12N1/21,C12N5/10,C12P21/08,C12Q1/68,G01N33/53,G01N33/PC  
574,  
PC C12N15/00,C12N5/00  
CC n=not determined(3243)  
FH Key Location/Qualifiers  
FT source 1..3283  
FT /organism="Homo sapiens (human)".  
FEATURES  
source 1..3283  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
ORIGIN  
Query Match 62.8%; Score 2143; DB 6; Length 3283;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1270 ACCCACTCCGGATACCTTCACGCTGTACCCCATCACAGTTGGCCCGTTCGCGCAT 1329  
DB 1141 ACCCACTCCGGATACCTTCACGCTGTACCCCATCACAGTTGGCCCGTTCGCGCAT 1200  
QY 1330 CATCACTCTTATCCAGAGCAGGAGTTGTAATCTCTTATCCCAACCCAGGCTGTGGC 1389  
DB 1201 CATCACTCTTATCCAGAGCAGGAGTTGTAATCTCTTATCCCAACCCAGGCTGTGGC 1260  
QY 1390 GCCATCATCGGGAAGAGGGGACACATCAAAACAGCTGGCGAGATTGCGCGGAGCCTCT 1449  
DB 1261 GCCATCATCGGGAAGAGGGGACACATCAAAACAGCTGGCGAGATTGCGCGGAGCCTCT 1320  
QY 1450 ATCAAGATTGCCCTCGGGAAGGCCAGAGCTGAGCGAAGAGATGGTATCATCAACGGG 1509  
DB 1321 ATCAAGATTGCCCTCGGGAAGGCCAGAGCTGAGCGAAGAGATGGTATCATCAACGGG 1380  
QY 1510 CCACCGGAAGCCAGTTCAAGGCCCCAGGACGGATCTTTGGGAACTGAAAGAGGAAAC 1569  
DB 1381 CCACCGGAAGCCAGTTCAAGGCCCCAGGACGGATCTTTGGGAACTGAAAGAGGAAAC 1440  
QY 1570 TTCTTTAAACCCCAAGGAAGTGAAGCTGGAAGCGCATATCAAGTGCCTCTTCCACA 1629  
DB 1441 TTCTTTAAACCCCAAGGAAGTGAAGCTGGAAGCGCATATCAAGTGCCTCTTCCACA 1500  
QY 1630 GCTGGCCGGGTGATTGGCAAGGTGGCAAGACCGTGAACCACTGCGAAGACTTAAACAGT 1689  
DB 1501 GCTGGCCGGGTGATTGGCAAGGTGGCAAGACCGTGAACCACTGCGAAGACTTAAACAGT 1560

QY 1690 GCAGAAAGTATCGTGCTCGTGAACCAACGCCAGATGAAAATGAGGAAGTGAATCGTCAGA 1749  
DB 1561 GCAGAAAGTATCGTGCTCGTGAACCAACGCCAGATGAAAATGAGGAAGTGAATCGTCAGA 1620  
QY 1750 ATTATCGGGCACCTTTCTTGTAGCCAGACTGACACAGCGAAGATCAGGGAATTTGTACAA 1809  
DB 1621 ATTATCGGGCACCTTTCTTGTAGCCAGACTGACACAGCGCAAGATCAGGNAATTTGTACAA 1680  
QY 1810 CAGGTGAAGCAGCAGGAGCAGAAATACCTTCAGGGAGTGCCTTCACAGGCGCAACAAGTGA 1869  
DB 1681 CAGGTGAAGCAGCAGGAGCAGAAATACCTTCAGGGAGTGCCTTCACAGGCGCAACAAGTGA 1740  
QY 1870 GGTCTCCACAGGCACGACGACCAACCAACGAGTGAATGTAGCCCTTCCAAACCTTCAGAGAA 1929  
DB 1741 GGTCTCCACAGGCACGACGACCAACCAACGAGTGAATGTAGCCCTTCCAAACCTTCAGAGAA 1800  
QY 1930 TGAGACCAAAACGACGACGACGATCGGAGCAAAACCAAGACCATCTGAGGAATGAGAA 1989  
DB 1801 TGAGACCAAAACGACGACGACGATCGGAGCAAAACCAAGACCATCTGAGGAATGAGAA 1860  
QY 1990 GTCTGCGAGGCGGCGCAGGGAATCTGCGCAGGCGCTTGAGAAACCCAGGGGCCGAGAGGG 2049  
DB 1861 GTCTGCGAGGCGGCGCAGGGAATCTGCGCAGGCGCTTGAGAAACCCAGGGGCCGAGAGGG 1920  
QY 2050 GCGGGAAGGTGACGACGAGTTTGCAGAAACCCAGGAGCCCGCTCCGCCCCCAGGGC 2109  
DB 1921 GCGGGAAGGTGACGACGAGTTTGCAGAAACCCAGGAGCCCGCTCCGCCCCCAGGGC 1980  
QY 2110 TTCTGCAAGGCTTCAGGCATCCACTTCACCATCCACTCGGATCTCTCGTAACCTCCACAGA 2169  
DB 1981 TTCTGCAAGGCTTCAGGCATCCACTTCACCATCCACTCGGATCTCTCGTAACCTCCACAGA 2040  
QY 2170 CGCTATCCCTTTTAGTTGAACATGATAGTGAACGTGTTCAAAGCCCAAGCAAAATGACAC 2229  
DB 2041 CGCTATCCCTTTTAGTTGAACATGATAGTGAACGTGTTCAAAGCCCAAGCAAAATGACAC 2100  
QY 2230 ACCCTTTTCTGCGCAAAATCGTCTGTGTACATGTGTGTACATATTAGAAAGGGAAGATG 2289  
DB 2101 ACCCTTTTCTGCGCAAAATCGTCTGTGTACATGTGTGTACATATTAGAAAGGGAAGATG 2160  
QY 2290 TTAAGATATGTGCTGTGGGTTACACAGGGTGCCTGCAGCGGTAATATATTTTAGAAAT 2349  
DB 2161 TTAAGATATGTGCTGTGGGTTACACAGGGTGCCTGCAGCGGTAATATATTTTAGAAAT 2220  
QY 2350 AATATATCAATAAATCAACTCAACTCCAAATTTTAAATCAATTAATTTTCTTTTCTTT 2409  
DB 2221 AATATATCAATAAATCAACTCAACTCCAAATTTTAAATCAATTAATTTTCTTTTCTTT 2280  
QY 2410 TTAAGAGAAAGCAGGCTTTCTAGACTTTTAAAGAAATAAAGTCTTTGGGAGGTCTCACGG 2469  
DB 2281 TTAAGAGAAAGCAGGCTTTCTAGACTTTTAAAGAAATAAAGTCTTTGGGAGGTCTCACGG 2340  
QY 2470 TGTAGAGAGAGCTTTGAGGCCACCCGCAAAAATTCACCCAGAGGAAATCTCGTCGA 2529  
DB 2341 TGTAGAGAGAGCTTTGAGGCCACCCGCAAAAATTCACCCAGAGGAAATCTCGTCGA 2400  
QY 2530 AGGACACTCAGCGAGTCTTGGATCAGCTGTGTATGTACACAGAGGATACCGTCTCTCT 2589  
DB 2401 AGGACACTCAGCGAGTCTTGGATCAGCTGTGTATGTACACAGAGGATACCGTCTCTCT 2460  
QY 2590 TGAAGAGGAAACTCTGTCACTCTCTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2649  
DB 2461 TGAAGAGGAAACTCTGTCACTCTCTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2520  
QY 2650 TTCAAGGTTTAAACTGGTTTTTGGATACCTGCTATATTAATTTCTGTCTCTCTCTCTCTGT 2709  
DB 2521 TTCAAGGTTTAAACTGGTTTTTGGATACCTGCTATATAAATTTCTGTCTCTCTCTCTCTGT 2580  
QY 2710 TATCTCTCCCTCCCTCCCTCCCTCTCTCTCCATCTCCATCTCTCTCTCTCTCTCTCTCTCAT 2769  
DB 2581 TATCTCTCCCTCCCTCCCTCCCTCTCTCTCCATCTCCATCTCTCTCTCTCTCTCTCTCTCAT 2640



QY	657	GGGGATCTCTGGTCC	CCACCCAGT	TTTGTGGTGC	CATCATCGGA	AAAGAGG	CGCTTG	ACCAT	716					
Db	634	GGGGATCTCTGGTCC	CCACCCAGT	TTTGTGGTGC	CATCATCGGA	AAAGAGG	CGCTTG	ACCAT	693					
QY	717	AAAGAAATCACTAA	AGCAGACCC	AGTCCCGGG	TAGATATCC	ATAGAAAA	GAGAACT	CTGG	776					
Db	694	AAAGAAATCACTAA	AGCAGACCC	AGTCCCGGG	TAGATATCC	ATAGAAAA	GAGAACT	CTGG	753					
QY	777	AGCTGAGAGAGCCT	GTACCATCC	ATGCCCAC	CCAGAGGG	GACCTTCT	GAAAGCAT	TCGG	836					
Db	754	AGCTGAGAGAGCCT	GTACCATCC	ATGCCCAC	CCAGAGGG	GACCTTCT	GAAAGCAT	TCGG	813					
QY	837	CATGATCTTTGAA	ATCATGCAG	AAAAAGG	CAGATGAG	CAAACTAG	CCGAAAG	ATTCC	896					
Db	814	CATGATCTTTGAA	ATCATGCAG	AAAAAGG	CAGATGAG	CAAACTAG	CCGAAAG	ATTCC	873					
QY	897	TCTGAAAAATCTT	GGGCACAA	TGCTTGGT	TGGAAG	ACTGAT	TGAAAA	AGAGG	956					
Db	874	TCTGAAAAATCTT	GGGCACAA	TGCTTGGT	TGGAAG	ACTGAT	TGAAAA	AGAGG	933					
QY	957	TTTGAAGAAAA	TTGAAATGA	ACATGA	ACGGG	ACCAAG	ATACAT	CTCATCTT	TTCAGAG	1016				
Db	934	TTTGAAGAAAA	TTGAAATGA	ACATGA	ACGGG	ACCAAG	ATACAT	CTCATCTT	TTCAGAG	993				
QY	1017	GAGCATATACA	CCCGGAA	AGAACCA	CTCACTGT	GAAAGG	CACAGT	TGAGCC	CTGTGCC	1076				
Db	994	GAGCATATACA	CCCGGAA	AGAACCA	CTCACTGT	GAAAGG	CACAGT	TGAGCC	CTGTGCC	1053				
QY	1077	TGCTGAGATAG	AGATTATGA	AGAGCTG	CGTGA	GGCCCTT	TGAAAA	TGATATG	CTGCTGT	1136				
Db	1054	TGCTGAGATAG	AGATTATGA	AGAGCTG	CGTGA	GGCCCTT	TGAAAA	TGATATG	CTGCTGT	1113				
QY	1137	TAACCAACAG	CAATCTG	ATCCAG	GGTTGA	ACCTC	CAGCG	CACTTGG	CACTCTT	CAAC	1196			
Db	1114	TAACCAACAG	CAATCTG	ATCCAG	GGTTGA	ACCTC	CAGCG	CACTTGG	CACTCTT	CAAC	1173			
QY	1197	AGGACTGTG	CGTCTATCT	CCACAG	CAGGG	CCCCCG	CGAGCT	CCCCCG	CTGCC	CCCTA	1256			
Db	1174	AGGACTGTG	CGTCTATCT	CCACAG	CAGGG	CCCCCG	CGAGCT	CCCCCG	CTGCC	CCCTA	1233			
QY	1257	CGACCCCTT	CACTAC	PACCACT	CGGAT	ACTTCT	CCAG	CTGTAT	CCCCCAT	CACAG	TTGG	1316		
Db	1234	CGACCCCTT	CACTAC	PACCACT	CGGAT	ACTTCT	CCAG	CTGTAT	CCCCCAT	CACAG	TTGG	1293		
QY	1317	CCGGTTCCG	CATCACT	CTTATCC	AGAGC	AGGAT	TGTAAT	CTCTT	CA	TCCCA	C	1376		
Db	1294	CCGGTTCCG	CATCACT	CTTATCC	AGAGC	AGGAT	TGTAAT	CTCTT	CA	TCCCA	C	1353		
QY	1377	CGAGGCTGTG	GGGCCAT	CATCGG	AGAGAG	GGGG	GCACAT	CAAA	CAGCTG	CGGAG	ATT	1436		
Db	1354	CGAGGCTGTG	GGGCCAT	CATCGG	AGAGAG	GGGG	GCACAT	CAAA	CAGCTG	CGGAG	ATT	1413		
QY	1437	CGCCGGAG	CCCTCTAT	CAAGATT	CGCCCTG	CGGAAG	GGCC	CAGAG	CTG	CAG	GGAAAG	GGT	1496	
Db	1414	CGCCGGAG	CCCTCTAT	CAAGATT	CGCCCTG	CGGAAG	GGCC	CAGAG	CTG	CAG	GGAAAG	GGT	1473	
QY	1497	CATCATCA	CGGG	CCACCG	GAAGCC	AGTTTCA	AGG	CCCGG	AGGAG	CGGAT	CTTTCGG	AAA	CT	1556
Db	1474	CATCATCA	CGGG	CCACCG	GAAGCC	AGTTTCA	AGG	CCCGG	AGGAG	CGGAT	CTTTCGG	AAA	CT	1533
QY	1557	GAAAGAGG	AAAACTT	CTTTAA	CCCCCA	AGAG	AGTGA	AGCTG	GGAAG	CGCAT	ATC	AG	1616	
Db	1534	GAAAGAGG	AAAACTT	CTTTAA	CCCCCA	AGAG	AGTGA	AGCTG	GGAAG	CGCAT	ATC	AG	1593	
QY	1617	GCCCTCTT	CCACAG	CTGG	CGGGT	GATTG	GC	CAAG	GTG	CA	ACCG	TGA	CT	1676
Db	1594	GCCCTCTT	CCACAG	CTGG	CGGGT	GATTG	GC	CAAG	GTG	CA	ACCG	TGA	CT	1653
QY	1677	GAACTTAA	CCAG	TGCAG	AAAGT	CTATCG	TG	CGCT	CTG	TG	ACCA	CCG	AG	1736
Db	1654	GAACTTAA	CCAG	TGCAG	AAAGT	CTATCG	TG	CGCT	CTG	TG	ACCA	CCG	AG	1713
QY	1737	AGTGAT	CTCAG	AAATTAT	CGGG	CACTTCTT	TGCT	AGCC	AGCTG	CA	CG	CG	CAAG	1796

Db	1714	AG	TGATCTG	TACAGAA	TTATTC	GGGCAC	TCTTTG	TAGCC	ACTG	CACAG	CGCAAG	ATCAG	1773
QY	1797	GG	AAATTTG	TACAAC	AGGTGA	AGCAG	CAGG	AGCAG	AGAA	TACCTC	CAGG	AGTCGC	1856
Db	1774	GG	AAATTTG	TACAAC	AGGTGA	AGCAG	CAGG	AGCAG	AGAA	TACCTC	CAGG	AGTCGC	1833
QY	1857	GC	CAGCAAG	TGAGG	TCC	CACAGG	CAAC	AGCA	AAAA	CAACG	ATGATG	AGCCCTT	1916
Db	1834	GC	CAGCAAG	TGAGG	TCC	CACAGG	CAAC	AGCA	AAAA	CAACG	ATGATG	AGCCCTT	1893
QY	1917	AC	ACTTGAC	AGAATG	AGAC	CAAA	CGCAG	CGCAG	ATCG	GGAG	CAAA	CCAAAG	1976
Db	1894	AC	ACTTGAC	AGAATG	AGAC	CAAA	CGCAG	CGCAG	ATCG	GGAG	CAAA	CCAAAG	1953
QY	1977	TG	AGGAATG	AGAAGT	CTG	CGGAG	CGCG	CCAGG	CACT	TGCG	AGG	CCCTGAG	2033
Db	1954	TG	AGGAATG	AGAAGT	CTG	CGGAG	CGCG	CCAGG	CACT	TGCG	AGG	CCCTGAG	2010
RESULT 9	AF057352												
LOCUS													
DEFINITION													
ACCESSION													
VERSION													
KEYWORDS													
SOURCE													
ORGANISM													
REFERENCE													
AUTHORS													
TITLE													
JOURNAL													
MEDLINE													
PUBMED													
REFERENCE													
AUTHORS													
TITLE													
JOURNAL													
FEATURES													
source													
gene													
CDS													
ORIGIN													
Query Match													
Best local Similarity													

53.0%; Score 1810; DB 9; Length 3667;  
99.9%; Pred. No. 0;

Matches 2100; Conservative 0; Mismatches 1; Indels 2; Gaps 2;			
QY	1270	ACCCATCCGGATACCTTCTCCAGCCTGTATCCCCCATCACAGATTTGGCCCGCTTCCCGCAT	1329
Db	1507	ACCCATCCGGATACCTTCTCCAGCCTGTATCCCCCATCACAGATTTGGCCCGCTTCCCGCAT	1566
QY	1330	CATCACTCTATCCAGAGCAGAGATGTGAATCTCTTATCCCAACCCAGCTGTGGGC	1389
Db	1567	CATCACTCTATCCAGAGCAGAGATGTGAATCTCTTATCCCAACCCAGCTGTGGGC	1626
QY	1390	GCATCATCGGGAAGAGGGGCACACATCAACACGCTGGCGAGATTCGCCGGAGCTCT	1449
Db	1627	GCCATCATCGGGAAGAGGGGCACACATCAACACGCTGGCGAGATTCGCCGGAGCTCT	1686
QY	1450	ATCAAGATTTGCCCTCCGGAAAGCCCGAGCGTCAGCGAAAGGATGTATCATCACCGGG	1509
Db	1687	ATCAAGATTTGCCCTCCGGAAAGCCCGAGCGTCAGCGAAAGGATGTATCATCACCGGG	1746
QY	1510	CCACCGGAAGCCAGTTCAAGGCCCGAGGACGGATCTTTGGGAACTGAAAGAGGAAAC	1569
Db	1747	CCACCGGAAGCCAGTTCAAGGCCCGAGGACGGATCTTTGGGAACTGAAAGAGGAAAC	1806
QY	1570	TTCTTTAAACCCAAAGAAAGTGAAGCTGGAAGCGCATATCAGAGTGCCTCTTCCACA	1629
Db	1807	TTCTTTAAACCCAAAGAAAGTGAAGCTGGAAGCGCATATCAGAGTGCCTCTTCCACA	1866
QY	1630	GCTGGCCGGTGAATGGCAAGAGTGCAAGA CCGTGAACGAACTGCAGAACTTAACCACT	1689
Db	1867	GCTGGCCGGTGAATGGCAAGAGTGCAAGA CCGTGAACGAACTGAACCACTTAACCACT	1926
QY	1690	GCAGAGTCATCGTGCCTCGTGACCAACCGCCAGATGCAAAATCAGAGAGTGCATCAG	1749
Db	1927	GCAGAGTCATCGTGCCTCGTGACCAACCGCCAGATGCAAAATCAGAGAGTGCATCAG	1986
QY	1750	ATTATCGGGCACTTTTGTAGCCAGACTGACAGCGCAAGATCAGGAAATTTGTACAA	1809
Db	1987	ATTATCGGGCACTTTTGTAGCCAGACTGACAGCGCAAGATCAGGAAATTTGTACAA	2046
QY	1810	CHAGGTGAAGCAGAGCAGAGAAATACCTCTAGGGAGTGCCTTCAAGCGCAGCAAGTGA	1869
Db	2047	CAGGTGAAGCAGCAGGAGCAGAAATACCTCTAGGGAGTGCCTTCAAGCGCAGCAAGTGA	2106
QY	1870	GGCTCCCAAGCAGCAGCAACCAACACAGATGATAGCCCTTCCAAACACCTGACAGAA	1929
Db	2107	GGCTCCCAAGCAGCAGCAACCAACACAGATGATAGCCCTTCCAAACACCTGACAGAA	2166
QY	1930	TCAGACCAAAACGAGCAGCCAGATCGGAGCAAAACCAAGAGCAATCTCAGGAAATGAGAA	1989
Db	2167	TGAGACCAAAACGAGCAGCCAGATCGGAGCAAAACCAAGAGCAATCTCAGGAAATGAGAA	2226
QY	1990	GTCTGGGAGGCGGCAGGGA CTCTGCGAGGCCCTTGAGAACCCAGGGCGGAGGAGGG	2049
Db	2227	GTCTGGGAGGCGGCAGGGA CTCTGCGAGGCCCTTGAGAACCCAGGGCGGAGGAGGG	2286
QY	2050	GGGGGAAGGTCAGCCAGGTTTGCAGAACCCACCGAGCCCGCTCCCGCCCGCCAGGGC	2109
Db	2287	GGGGGAAGGTCAGCCAGGTTTGCAGAACCCACCGAGCCCGCTCCCGCCCGCCAGGGC	2346
QY	2110	TTCTCGAGGCTTCAGCCATCCACTTCACCATCCACTCGGATCTCTCTGAACTCCCAAGA	2169
Db	2347	TTCTCGAGGCTTCAGCCATCCACTTCACCATCCACTCGGATCTCTCTGAACTCCCAAGA	2406
QY	2170	CGTATCCCTTTTGTAGTTGAACATAAGGTGAACCGTGTTCAAAGCCAAGCAAAATGCAC	2229
Db	2407	CGTATCCCTTTTGTAGTTGAACATAAGGTGAACCGTGTTCAAAGCCAAGCAAAATGCAC	2466
QY	2230	ACCTTTTCTGTGGCAATCGTCTGTACATGTGTGTACATATTTAGAAAGGAGATG	2289
Db	2467	ACCTTTTCTGTGGCAATCGTCTGTGTACATGTGTGTACATATTTAGAAAGGAGATG	2526
QY	2290	TTAAGATATGTGGCCTGTGGTTTACACAGGGTGCCTGCGAGCGGTATATATTTAGAAAT	2349
Db	2527	TTAAGATATGTGGCCTGTGGTTTACACAGGGTGCCTGCGAGCGGTATATATTTAGAAAT	2586

QY	2350	AAATATACAAATAACTCAACTCAACTCAATTTTTTAATCAATTAATTAATTTTTTCTTT	2409
Db	2587	AAATATACAAATAACTCAACTCAACTCAATTTTTTAATCAATTAATTAATTTTTTCTTT	2646
QY	2410	TTAAAGAGAAACAGAGCTTTCTAGACTTTAAAGATAAAGCTTTTGGAGGCTTCACGG	2469
Db	2647	TTAAAGAGAAACAGAGCTTTCTAGACTTTAAAGATAAAGCTTTTGGAGGCTTCACGG	2706
QY	2470	TGTAGAGAGAGCTTTGAGGCCACCGCACAAAATTCACCCAGAGGAAATCTCGTCGA	2529
Db	2707	TGTAGAGAGAGCTTTGAGGCCACCGCACAAAATTCACCCAGAGGAAATCTCGTCGA	2766
QY	2530	AGGACATCACCGCAGTTCTCGATCACCTGTGTATGTCAACAGAGGGATA CCGTCTCCT	2589
Db	2767	AGGACATCACCGCAGTTCTCGATCACCTGTGTATGTCAACAGAGGGATA CCGTCTCCT	2826
QY	2590	TGAAGAGGAACTGTGTCACTCCTCATGCTCTAGCTCATACACCCATTTCTCTTGC	2649
Db	2827	TGAAGAGGAACTGTGTCACTCCTCATGCTCTAGCTCATACACCCATTTCTCTTGC	2886
QY	2650	TTTCAAGGTTTTAAACTGGTTTTTTTGCACTACTGCTATATAAATCTCTGTCTCTCTGTT	2709
Db	2887	TTTCAAGGTTTTAAACTGGTTTTTTTGCACTACTGCTATATAAATCTCTGTCTCTCTGTT	2946
QY	2710	TATCTCTCCCTCCCTCCCTCCCTCTTCTCCATCTCCATTTCTTCAATTTCTCAT	2769
Db	2947	TATCTCTCCCTCCCTCCCTCCCTCTTCTCCATCTCCATTTCTTCAATTTCTCAT	3006
QY	2770	CCCTCCATCTCAATCCCGTATCTAGCACCCGCCCGCCCGCCCGCCCGCCCGCTCA	2829
Db	3007	CCCTCCATCTCAATCCCGTATCTAGCGCA - CCGCCCGCCCGCCCGCCCGCCCGCTCA	3065
QY	2830	GTATCACATFACACAAAAGGAAACAAAGCGAAACACAAAACAGCCTCAAATTACACTT	2889
Db	3066	GTATCACATFACACAAAAGGAAACAAAGCGAAACACAAAACAGCCTCAAATTACACTT	3125
QY	2890	GGTTACTCAAAAGAACAGAGTCAATGGTACTTGTCTTGTAGCGTTTGGAGGAGGAAAA	2949
Db	3126	GGTTACTCAAAAGAACAGAGTCAATGGTACTTGTCTTGTAGCGTTTGGAGGAGGAAAA	3185
QY	2950	GGAAACCCACCAACCAACCAATCAACCAACCAACCAACCAACCAACCAACCAACCA	3009
Db	3186	GGAAACCCACCAACCAACCAATCAACCAACCAACCAACCAACCAACCAACCAACCA	3245
QY	3010	ATTTTGTCTTTTGTGATTTTGTGTATAGCCATCAATATTCAGCAAAATGATTCCTTTC	3069
Db	3246	ATTTTGTCTTTTGTGATTTTGTGTATAGCCATCAATATTCAGCAAAATGATTCCTTTC	3305
QY	3070	TTTT - AAAAAAAAATGTGGAGGAAAGTAGAAATTTTCAAGGTTTGGCCCGAGGCGT	3128
Db	3306	TTTTAAAAAAAATGTGGAGGAAAGTAGAAATTTTCAAGGTTTGGCCCGAGGCGT	3365
QY	3129	TAAATTCAGAGATTTTAAACGAGAAACACACAGAGAGAGCTACCTCAGTGTTTTT	3188
Db	3366	TAAATTCAGAGATTTTAAACGAGAAACACACAGAGAGAGCTACCTCAGTGTTTTT	3425
QY	3189	ACCTCAGCACCTTGTCTTGTCTTTCCCTTAGAGATTTTGTAAAGCTGATTTGGAGCA	3248
Db	3426	ACCTCAGCACCTTGTCTTGTCTTTCCCTTAGAGATTTTGTAAAGCTGATTTGGAGCA	3485
QY	3249	TTTTTTTATTTTTTAAATAAAATGAGTTGGAAAAAATAAGATATCACTGCCAGCCT	3308
Db	3486	TTTTTTTATTTTTTAAATAAAATGAGTTGGAAAAAATAAGATATCACTGCCAGCCT	3545
QY	3309	GGAGAGGTGACAGTCCAGTGTGCAACAGCTTCTGAAATGTCTTCCGTAGCCAGA	3368
Db	3546	GGAGAGGTGACAGTCCAGTGTGCAACAGCTTCTTGAATTTGTCTTCCGTAGCCAGA	3605
QY	3369	ACC 3371	
Db	3606	ACC 3608	











QY	2914	ATGGTACTTGTCTAGCGTTTGGAGAGGAAAAACAGGAGACCACCAACCAACCAATCA	2973
Db	26574	ATGGTACTTGTCTAGCGTTTGGAGAGGAAAAACAGGAGACCACCAACCAACCAATCA	26515
QY	2974	ACCAACCAAGAAAAAATCCACATGAAGAATGATTTCTCTTTTGGCAATTTGGTG	3033
Db	26514	ACCAACCAAGAAAAAATCCACATGAAGAATGATTTCTCTTTTGGCAATTTGGTG	26455
QY	3034	TATAGCCATCAATATTCAGCAAAATGATTCCTTTT-AAAAAATAAATGTGGAGA	3092
Db	26454	TATAGCCATCAATATTCAGCAAAATGATTCCTTTT-AAAAAATAAATGTGGAGA	26395
QY	3093	AGTGAATTTACCAAGGTTGTGCCCGAGGCGTTAAATTCACAGATTTTTTAAACA	3152
Db	26394	AGTGAATTTACCAAGGTTGTGCCCGAGGCGTTAAATTCACAGATTTTTTAAACA	26335
QY	3153	GAAAAACACACAGAGAAGCTACCTCAGGTGTTTTTACCTCAGCAGCTTGTCTTGTGT	3212
Db	26334	GAAAAACACACAGAGAAGCTACCTCAGGTGTTTTTACCTCAGCAGCTTGTCTTGTGT	26275
QY	3213	TCCTTAGAGATTTGTAAGCTGATGATGTTGAGCATTTTTTTTATTTTAAATAAAT	3272
Db	26274	TCCTTAGAGATTTGTAAGCTGATGATGTTGAGCATTTTTTTTATTTTAAATAAAT	26215
QY	3273	GAGTTGGAATAAATAAGATATCACTGCCAGCTCGAGAGTGCACAGTCCAGTGTG	3332
Db	26214	GAGTTGGAATAAATAAGATATCACTGCCAGCTCGAGAGTGCACAGTCCAGTGTG	26155
QY	3333	CAACAGCTTCTGTAATGTTCTTCGCTAGCCAGCAAGAAC	3371
Db	26154	CAACAGCTTCTGTAATGTTCTTCGCTAGCCAGCAAGAAC	26116
RESULT 12	AC134942/c		
LOCUS	AC134942	208273 bp	DNA linear HTG 27-MAR-2003
DEFINITION	Homo sapiens chromosome 3 clone RP11-164H23, WORKING DRAFT SEQUENCE, 21 unordered pieces.		
ACCESSION	AC134942		
VERSION	AC134942.3	GI:292933991	
KEYWORDS	HTG; HTGS PHASE1; HTGS_DRAFT; HTGS_FULLTOP.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-osman, F.R., Allen, C., Alsbrooks, S.L., Amaratunge, H.C., Are, J.R., Ayele, M., Banks, T., Barbara, J., Benton, J., Binage, K., Blankenburg, K., Bonnin, D., Bouck, J., Bowie, S., Brieva, M., Brown, E., Brown, N., Bryant, N.P., Buay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C., Carron, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C., Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R., Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A., Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H., Douthwaite, K.J., Draper, H., Dugan-Kocha, S., Durbin, K.J., Earhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M., Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P., Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R., Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K., Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J., Hernandez, O., Hodgson, A., Hogue, M., Holloway, C., Hollins, B., Jacobson, B., Jia, X., Johnson, R., Jolivet, S., Joudah, S., Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C., Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L., Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W., Loubege, H., Lozado, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J., Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E., Massey, B., Mawhney, E., McLeod, M.P., Meador, M., Mei, G., Metzker, M., Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S., Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N.,		

Nguyen, N., Nickerson, E., Nwokenkwo, S., Ogih, M., Okwuonu, G., Oragunye, N., Oviado, R., Pace, A., Payton, B., Peery, J., Perez, I., Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y., Rivers, M., Rojas, A., Rojokan, I., Rolfe, M., Ruiz, S., Savery, G., Scherer, S., Scott, G., Shen, H., Shooshtari, N., Sisson, I., Sodergren, E., Sonaike, T., Sparks, A., Stanley, H., Stone, H., Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H., Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S., Umani, K., Vasquez, L., Vera, V., Villalon, D., Vinson, R., Wang, Q., Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S., Williams, G., Williamson, A., Wleczek, R., Wooden, S., Worley, K., Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D., Weinstein, G., and Gibbs, R.	
Direct Submission	
Unpublished	
2 (bases 1 to 208273)	
Worley, K.C.	
Direct Submission	
Submitted (03-OCT-2002)	Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 208273)	
Worley, K.C.	
Direct Submission	
Submitted (27-MAR-2003)	Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On Mar 27, 2003	this sequence version replaced gi:27819431.
Center: Baylor College of Medicine	
Center code: BCM	
Web site: <a href="http://www.hgsc.bcm.tmc.edu/">http://www.hgsc.bcm.tmc.edu/</a>	
Drafting Center Code: BCM	
Contact: hgsc-help@bcm.tmc.edu	
----- Project Information	
Center project name: HDIO	
Center clone name: RP11-164H23	
----- Summary Statistics	
Sequencing vector: Plasmid;	
Chemistry: Dye-terminator Big Dye; 100% of reads	
Assembly program: Phrap; version 0.990329	
Consensus quality: 203806 bases at least Q40	
Consensus quality: 206019 bases at least Q30	
Consensus quality: 207337 bases at least Q20	
Estimated insert size: 200535; sum-of-contigs estimation	
Quality coverage: 10x in Q20 bases; sum-of-contigs estimation	
-----	
* NOTE: Estimated insert size may differ from sequence length	
* (see <a href="http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html">http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html</a> ).	
* NOTE: This is a 'working draft' sequence. It currently	
* consists of 21 contigs. The true order of the pieces	
* is not known and their order in this sequence record is	
* arbitrary. Gaps between the contigs are represented as	
* runs of N, but the exact sizes of the gaps are unknown.	
* This record will be updated with the finished sequence	
* as soon as it is available and the accession number will	
* be preserved.	
*	1 1288: contig of 1288 bp in length
*	1289 1388: gap of unknown length
*	1389 2750: contig of 1362 bp in length
*	2751 2850: gap of unknown length
*	2851 4127: contig of 1277 bp in length
*	4128 4227: gap of unknown length
*	4228 5868: contig of 1641 bp in length
*	5869 5969: gap of unknown length
*	5969 7223: contig of 1255 bp in length
*	7224 7323: gap of unknown length
*	7324 8668: contig of 1345 bp in length
*	8669 10070: contig of 1302 bp in length
*	10071 11650: contig of 1480 bp in length
*	11651 11750: gap of unknown length

\* 11751 13080: contig of 1330 bp in length  
\* 13081 13180: gap of unknown length  
\* 13181 14705: contig of 1525 bp in length  
\* 14706 14805: gap of unknown length  
\* 14806 15952: contig of 1147 bp in length  
\* 15953 16052: gap of unknown length  
\* 16053 17790: contig of 1738 bp in length  
\* 17791 17890: gap of unknown length  
\* 17891 20675: contig of 2785 bp in length  
\* 20676 20775: gap of unknown length  
\* 20776 24752: contig of 3977 bp in length  
\* 24753 24852: gap of unknown length  
\* 24853 35887: contig of 11035 bp in length  
\* 35888 35987: gap of unknown length  
\* 35988 51687: contig of 15700 bp in length  
\* 51688 51788: gap of unknown length  
\* 51788 67942: contig of 16155 bp in length  
\* 67943 68042: gap of unknown length  
\* 68043 99019: contig of 30977 bp in length  
\* 99020 99119: gap of unknown length  
\* 99120 132270: contig of 33151 bp in length  
\* 132271 132370: gap of unknown length  
\* 132371 16641: contig of 34271 bp in length  
\* 16642 16674: gap of unknown length  
\* 16674 208273: contig of 41532 bp in length.

## FEATURES

source  
1..208273  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="3"  
/clone="RP11-164H23"

## ORIGIN

Query Match 38.0%; Score 1296; DB 2; Length 208273;  
Best local Similarity 99.8%; Pred. No. 0;  
Matches 1596; Conservative 0; Mismatches 0; Indels 3; Gaps 2;  
  
QY 1774 CAGACTGCAGCGCAAGATCAGGAAATTGTACAAAGGTGAAGCAGCAGGACAGAAA 1833  
DB 82061 CAGACTGCAGCGCAAGATCAGGAAATTGTACAAAGGTGAAGCAGCAGGACAGAAA 82002  
  
QY 1834 TACCTCAGGAGTGGCTCAGCGCAGCAAGTGAAGCTCCACAGGCACACGCAAAAC 1893  
DB 82001 TACCTCAGGAGTGGCTCAGCGCAGCAAGTGAAGCTCCACAGGCACACGCAAAAC 81942  
  
QY 1894 AACGGATGATGTAGCCCTTCCACACTGCACAGAAATGAGACCAACGCGAGCCAGCAGA 1953  
DB 81941 AACGGATGATGTAGCCCTTCCACACTGCACAGAAATGAGACCAACGCGAGCCAGCAGA 81882  
  
QY 1954 TCGGGAGCAACCAAGAACCATCTGAGGAATGAGAAGTCTGCGGAGGCGCCAGGGACTC 2013  
DB 81881 TCGGGAGCAACCAAGAACCATCTGAGGAATGAGAAGTCTGCGGAGGCGCCAGGGACTC 81822  
  
QY 2014 TCGCGAGGCCCTGAGAAACCCAGGGCCGAGGCGCGGGAAGGTGACGCCAGTTTGC 2073  
DB 81821 TCGCGAGGCCCTGAGAAACCCAGGGCCGAGGCGCGGGAAGGTGACGCCAGTTTGC 81762  
  
QY 2074 CAGAAACACCGAGCCCGCTCCGCCGCCAGGCTTCTGACGGCTTCAGCCATCCACT 2133  
DB 81761 CAGAAACACCGAGCCCGCTCCGCCGCCAGGCTTCTGACGGCTTCAGCCATCCACT 81702  
  
QY 2134 TCACCATCCACTCGGATCTCTCTGAACCTCCACGAGCTATCCCTTTTAGTTGAACATA 2193  
DB 81701 TCACCATCCACTCGGATCTCTCTGAACCTCCACGAGCTATCCCTTTTAGTTGAACATA 81642  
  
QY 2194 CATAGGTGAACGTGTTCAAGACCAAGCAAAATGACACCCCTTTTCTGTGGGCAATCGTC 2253  
DB 81641 CATAGGTGAACGTGTTCAAGACCAAGCAAAATGACACCCCTTTTCTGTGGGCAATCGTC 81582  
  
QY 2254 TCTGTACATGTGTGTACATATTAGAAAGGAAGATGTTAAGATATGCGCCTGTGGGTTA 2313  
DB 81581 TCTGTACATGTGTGTGTACATATTAGAAAGGAAGATGTTAAGATATGCGCCTGTGGGTTA 81522

QY 2314 CACAGGGTGCTGCAGCGGTAAATATATATTTTAGAAATATATATCAAAATCACTCAACTAAC 2373  
DB 81521 CACAGGGTGCTGCAGCGGTAAATATATTTTAGAAATATATATCAAAATCACTCAACTAAC 81462  
  
QY 2374 TCCAAATTTTAAATCAATATTAATTTTCTTTTAAAGAGAAAGCAGGCTTTTCTA 2433  
DB 81461 TCCAAATTTTAAATCAATATTAATTTTCTTTTAAAGAGAAAGCAGGCTTTTCTA 81402  
  
QY 2434 GACTTTAAAGAAATAAAGTCTTTGGAGGTCTCACGGGTGAGAGGAGCTTTTGAGGCCAC 2493  
DB 81401 GACTTTAAAGAAATAAAGTCTTTGGAGGTCTCACGGGTGAGAGGAGCTTTTGAGGCCAC 81342  
  
QY 2494 CCGCACAATAATTCACCCAGAGGGAATCTCGTGGAGAGGACATCACGGCAGTCTGAT 2553  
DB 81341 CCGCACAATAATTCACCCAGAGGGAATCTCGTGGAGAGGACATCACGGCAGTCTGAT 81282  
  
QY 2554 CACTGTGTATGTCAACAGAGGAGATACCGTCTCTTTGAAGAGGAAACTCTGTCACTCCT 2613  
DB 81281 CACTGTGTATGTCAACAGAGGAGATACCGTCTCTTTGAAGAGGAAACTCTGTCACTCCT 81222  
  
QY 2614 CATGCTGTCTAGCTCATACACCAATTTCTTTGCTTCTCAGGTTTAAACTGGTTTT 2673  
DB 81221 CATGCTGTCTAGCTCATACACCAATTTCTTTGCTTCTCAGGTTTAAACTGGTTTT 81162  
  
QY 2674 TGCATACGTATATTAATTTCTGTCTCTCTCTGTTTATCTCTCCCTCCCTCCCTCCC 2733  
DB 81161 TGCATACGTATATTAATTTCTGTCTCTCTCTGTTTATCTCTCCCTCCCTCCCTCCC 81102  
  
QY 2734 CTCTCTCTCCATCTCCATTTCTTTGAATTTCTCTCATCCCTCCATCTCAATCCCGTATCTA 2793  
DB 81101 CTCTCTCTCCATCTCCATTTCTTTGAATTTCTCTCATCCCTCCATCTCAATCCCGTATCTA 81042  
  
QY 2794 CGCACCCCCCCCCCGAGCAAGCAGTCTCTGAGTATCATATCACACAAAGGAACA 2853  
DB 81041 CGCA--CCCCCCCCCGAGCAAGCAGTCTCTGAGTATCATATCACACAAAGGAACA 80984  
  
QY 2854 AAAGCGAAACACACAAACAGCCTCACTTACACTTGGTTTACTCAAAGAACAGAGTCA 2913  
DB 80983 AAAGCGAAACACACAAACAGCCTCACTTACACTTGGTTTACTCAAAGAACAGAGTCA 80924  
  
QY 2914 ATGTACTTGTCTCTAGCGTTTGGAGAGAAACAGGAAACCCACCAACCAACCAATCA 2973  
DB 80923 ATGTACTTGTCTCTAGCGTTTGGAGAGAAACAGGAAACCCACCAACCAATCA 80864  
  
QY 2974 ACCAAACAAAGAAAAATTCACAATGAAGAATGATTTTGTCTTTTGTGATTTTGGTG 3033  
DB 80863 ACCAAACAAAGAAAAATTCACAATGAAGAATGATTTTGTCTTTTGTGATTTTGGTG 80804  
  
QY 3034 TATAAGCCATCAATATTCAGCAAAATGATTCCTTTCTTTT--AAAAAATAATGTGGAGGA 3092  
DB 80803 TATAAGCCATCAATATTCAGCAAAATGATTCCTTTCTTTTAAATAATAATAATGTGGAGGA 80744  
  
QY 3093 AAGTAGAAATTTACCAAGTTTGTGGCCAGGCGTTAAATTCACAGATTTTTTAAACGA 3152  
DB 80743 AAGTAGAAATTTACCAAGTTTGTGGCCAGGCGTTAAATTCACAGATTTTTTAAACGA 80684  
  
QY 3153 GAAAAACACACAGAGAGTACTCCTCAGGTGTTTTTACCTCAGACCTTGCTCTTGTT 3212  
DB 80683 GAAAAACACACAGAGAGTACTCCTCAGGTGTTTTTACCTCAGACCTTGCTCTTGTT 80624  
  
QY 3213 TCCCTTAGAGATTTTGTAAAGCTGATGTGGAGCATTTTTTATTTTAAATAAATAAT 3272  
DB 80623 TCCCTTAGAGATTTTGTAAAGCTGATGTGGAGCATTTTTTATTTTAAATAAATAAT 80564  
  
QY 3273 GAGTTGGAAAAAATAAGATATCACTGCCAGCCTGGAGAGGTCAGTCCAGAGTGT 3332  
DB 80563 GAGTTGGAAAAAATAAGATATCACTGCCAGCCTGGAGAGGTCAGTCCAGAGTGT 80504  
  
QY 3333 CACAGCTGTCTGAATTTGTCTTCGCGTACCCAGAAC 3371  
DB 80503 CACAGCTGTCTGAATTTGTCTTCGCGTACCCAGAAC 80465

```

RESULT 13
AC126373/c
LOCUS
DEFINITION Homo sapiens chromosome 18 clone CTD-2307I13 map 18, LOW-PASS
SEQUENCE SAMPLING.
AC126373
ACCESSION AC126373
VERSION AC126373.1 GI:21699324
KEYWORDS HTG; HTGS PHASE0.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 50320)
AUTHORS Birren,B., Nusbaum,C. and Lander,E.
TITLE Homo sapiens chromosome 18, clone CTD-2307I13
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 50320)
AUTHORS Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L., Boukhgalter,B.,
Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J.,
Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B.,
Horton,L., Hulme,W., Illiev,I., Johnson,R., Jones,C., Kamat,A.,
Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K.,
Liu,G., Maclean,C., Macdonald,P., Major,J., Matthews,C.,
McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,
O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,
Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P.,
Roman,J., Roy,A., Schauer,S., Schupback,R., Seaman,S., Severy,P.,
Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J.,
Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,
Zembek,L., Zimmer,A. and Zody,M.
DIRECT SUBMISSION
TITLE Submitted (05-JUL-2002) Whitehead Institute/MIT Center for Genome
JOURNAL Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L27661
Center clone name: 2307_I_13
-----
* NOTE: This record contains 59 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.
1 766: contig of 766 bp in length
* 767 866: gap of 100 bp
* 867 1574: contig of 708 bp in length
* 1575 1674: gap of 100 bp
* 1675 2412: contig of 738 bp in length
* 2413 2512: gap of 100 bp
* 2513 3283: contig of 771 bp in length
* 3284 3384: gap of 100 bp
* 3384 4124: contig of 741 bp in length
* 4125 4224: gap of 100 bp
* 4225 4938: contig of 714 bp in length
*
5038: gap of 100 bp
5820: contig of 782 bp in length
5921 5920: gap of 100 bp
6685: contig of 765 bp in length
6785: gap of 100 bp
6786 7580: contig of 795 bp in length
7581 7680: gap of 100 bp
7681 8432: contig of 752 bp in length
8433 8532: gap of 100 bp
8533 9281: contig of 749 bp in length
9282 9381: gap of 100 bp
9382 10136: contig of 755 bp in length
10137 10236: gap of 100 bp
10237 10967: contig of 731 bp in length
11068 11787: contig of 720 bp in length
11788 11887: gap of 100 bp
11888 12655: contig of 768 bp in length
12656 12755: gap of 100 bp
12756 13497: contig of 742 bp in length
13498 13597: gap of 100 bp
13598 14365: contig of 768 bp in length
14366 14465: gap of 100 bp
14466 15203: contig of 738 bp in length
15204 15303: gap of 100 bp
15304 16010: contig of 707 bp in length
16011 16110: gap of 100 bp
16111 16872: contig of 762 bp in length
16873 16972: gap of 100 bp
16973 17715: contig of 743 bp in length
17716 17815: gap of 100 bp
17816 18587: contig of 772 bp in length
18588 18687: gap of 100 bp
18688 19464: contig of 777 bp in length
19465 19564: gap of 100 bp
19565 20303: contig of 739 bp in length
20304 20403: gap of 100 bp
20404 21135: contig of 732 bp in length
21136 21235: gap of 100 bp
21236 21999: contig of 764 bp in length
22000 22099: gap of 100 bp
22099 22846: contig of 747 bp in length
22847 22946: gap of 100 bp
22947 23709: contig of 763 bp in length
23710 23809: gap of 100 bp
23809 24525: contig of 716 bp in length
24526 24625: gap of 100 bp
24626 25371: contig of 746 bp in length
25372 25471: gap of 100 bp
25472 26239: contig of 768 bp in length
26240 26339: gap of 100 bp
26340 27090: contig of 751 bp in length
27091 27190: gap of 100 bp
27191 27946: contig of 756 bp in length
27947 28046: gap of 100 bp
28047 28808: contig of 762 bp in length
28809 28908: gap of 100 bp
28909 29669: contig of 761 bp in length
29670 29769: gap of 100 bp
29770 30550: contig of 781 bp in length
30551 30650: gap of 100 bp
30651 31423: contig of 773 bp in length
31424 31523: gap of 100 bp
31524 32279: contig of 756 bp in length
32280 32379: gap of 100 bp
32380 33129: contig of 750 bp in length
33130 33229: gap of 100 bp
33230 33984: contig of 755 bp in length
33985 34084: gap of 100 bp
34085 34828: contig of 744 bp in length
34829 34929: gap of 100 bp
34930 35665: contig of 737 bp in length
35666 35765: gap of 100 bp

```

35766 36535: contig of 770 bp in length  
36536 36536: gap of 100 bp  
36536 37412: contig of 777 bp in length  
37413 37512: gap of 100 bp  
37513 38284: contig of 772 bp in length  
38285 38384: gap of 100 bp  
38385 39173: contig of 789 bp in length  
39174 39273: gap of 100 bp  
39274 40051: contig of 778 bp in length  
40052 40151: gap of 100 bp  
40152 40900: contig of 749 bp in length  
40901 41000: gap of 100 bp  
41001 41778: contig of 778 bp in length  
41779 41878: gap of 100 bp  
41879 42645: contig of 767 bp in length  
42646 42745: gap of 100 bp  
42746 43454: contig of 709 bp in length  
43455 43554: gap of 100 bp  
43555 44314: contig of 760 bp in length  
44315 44414: gap of 100 bp  
44415 45158: contig of 744 bp in length  
45159 45258: gap of 100 bp  
45259 46040: contig of 782 bp in length  
46041 46140: gap of 100 bp  
46141 46894: contig of 754 bp in length  
46895 46994: gap of 100 bp  
46995 47754: contig of 760 bp in length  
47755 47854: gap of 100 bp  
47855 48596: contig of 742 bp in length  
48597 48696: gap of 100 bp  
48697 49446: contig of 750 bp in length  
49447 49546: gap of 100 bp  
49547 50320: contig of 774 bp in length.

FEATURES  
source  
1..50320  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="18"  
/map="18"  
/clone="CTD-2307I13"  
/clone\_lib="CITD1 Human BAC"

ORIGIN  
Query Match 17.2%; Score 586; DB 2; Length 50320;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 586; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2205 GTGTTCAAGCCAAAGCAAAATGCACACCCCTTTTCTGTGGCAAAATCGTCTGTACATGT 2264  
DB 10940 GTGTTCAAGCCAAAGCAAAATGCACACCCCTTTTCTGTGGCAAAATCGTCTGTACATGT 10881

QY 2265 GTGTACATATAGAAAGGAGATGTTAAGATATGCGCTGTGGTTACACAGGGTGC 2324  
DB 10880 GTGTACATATAGAAAGGAGATGTTAAGATATGCGCTGTGGTTACACAGGGTGC 10821

QY 2325 TGCAGCGGTATATATTTTGAATAATATATCAATAACTCAACTAACTCCAAATTTTA 2384  
DB 10820 TGCAGCGGTATATATTTTGAATAATATATCAATAACTCAACTAACTCCAAATTTTA 10761

QY 2385 ATCAATATTAATTTTTTTTCTTTTAAAGAGAAAGCAGCGTTTCTAGACATTTAAAGA 2444  
DB 10760 ATCAATATTAATTTTTTTTCTTTTAAAGAGAAAGCAGCGTTTCTAGACATTTAAAGA 10701

QY 2445 ATAAAGTCTTTGGAGTCTCAGGTGTAGAGAGGCTTTGAGGCCACCGCACAAAT 2504  
DB 10700 ATAAAGTCTTTGGAGTCTCAGGTGTAGAGAGGCTTTGAGGCCACCGCACAAAT 10641

QY 2505 TCACCCAGAGGGAATCTCGTCGGAAGGACACTCAGCGGAGTTCTGGATCACCTGTGTAT 2564  
DB 10640 TCACCCAGAGGGAATCTCGTCGGAAGGACACTCAGCGGAGTTCTGGATCACCTGTGTAT 10581

QY 2565 GTCACAGAGAGGATACCGTCTCTCTTGAAGAGAAACTCTGTCACTCTCATCGCTGTCT 2624

DB 10580 GTCAACAGAGGGATACCGTCTCTTGAAGAGAAACTCTGTCACTCTCTCATGCTGCT 10521

QY 2625 AGCTCATACACCCCATTTCTCTTTGCTTTCACAGAGTTTAAACTGGTTTTCATACACTGCT 2684

DB 10520 AGCTCATACACCCCATTTCTCTTTGCTTTCACAGAGTTTAAACTGGTTTTCATACACTGCT 10461

QY 2685 ATATAATCTCTGCTCTCTCTCTGTTTATCTCTCCCTCCCTCCCTCCCTCCCTTCTTCTCCA 2744

DB 10460 ATATAATCTCTGCTCTCTCTCTGTTTATCTCTCCCTCCCTCCCTCCCTTCTTCTCCA 10401

QY 2745 TCTCCATCTTTTGAATTTCTCATCCCTCATCTCAATCCCGTAT 2790

DB 10400 TCTCCATCTTTTGAATTTCTCATCCCTCATCTCAATCCCGTAT 10355

RESULT 14  
BD221443

LOCUS BD221443 710 bp DNA linear PAT 17-JUL-2003

DEFINITION Human gene and gene expression product V.

ACCESSION BD221443

VERSION BD221443.1 GI:33031213

KEYWORDS JP 2002534055-A/2556.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
Williams, L.T., Escobedo, J., Innis, M.A., Garcia, P.D., Klinger, J.S.,  
Reinhard, C., Giese, K., Randazzo, F., Kennedy, G.C., Pot, D.,  
Kassam, A., Lamson, G., Drmanac, R., Crkvenjakov, R., Dickson, M.,  
Drmanac, S., Labat, I., Leshkowitz, D., Kita, D., Garcia, V., Jones, L.W.  
and Crain, B.S.

REFERENCE  
1. (bases 1 to 710)  
Innis, M.A., Williams, L.T., Escobedo, J., Garcia, P.D., Klinger, J.S.,  
Reinhard, C., Giese, K., Randazzo, F., Kennedy, G.C., Pot, D.,  
Kassam, A., Lamson, G., Drmanac, R., Crkvenjakov, R., Dickson, M.,  
Drmanac, S., Labat, I., Leshkowitz, D., Kita, D., Garcia, V., Jones, L.W.  
and Crain, B.S.

AUTHORS

TITLE Human gene and gene expression product V

JOURNAL Patent: JP 2002534055-A 2556 15-OCT-2002;  
CHIRON CORP, HYSEQ INC

COMMENT OS Homo sapiens (human)  
PN JP 2002534055-A/2556  
PD 15-OCT-2002  
PF 13-MAY-1999 JP 2000548466  
PR 14-MAY-1998 US 60/085426, 15-MAY-1998 US 60/085537 PR  
15-MAY-1998 US 60/085696, 21-OCT-1998 US 60/105234 PR  
27-OCT-1998 US 60/105877

PI LOUIS T WILLIAMS, JAIME ESCOBEDO, MICHAEL A INNIS, PABLO PI  
DOMINGUEZ GARCIA,  
PI JULIE SUDDUTH KLINGER, CHRISTOPH REINHARD, KLAUSE GIESE, FILIPPO  
PI RANDAZZO,  
PI GIULIA C KENNEDY, DAVID POT, ALTAF KASSAM, GEORGE LAMSON, RADOJE  
PI DRMANAC,  
PI RADOMIR CRKVENJAKOV, MARK DICKSON, SNEZANA DRMANAC, IVAN LABAT,  
PI DENA LESHKOWITZ, DAVID KITA, VERONICA GARCIA, LEE WILLIAM JONES,  
PI BIRGIT STACHE CRAIN  
PC C12N15/00, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21 PC  
C12N5/00, C12Q1/68,  
PC C12N15/00, C12N5/00  
n = A, T, C or G  
CC Key Location/Qualifiers  
FH misc location (1)..(710).  
FT Location/Qualifiers  
source  
1..710  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"

ORIGIN  
Query Match 16.8%; Score 567; DB 6; Length 710;  
Best Local Similarity 100.0%; Pred. No. 2.5e-309;  
Matches 567; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 534 CCGGATGAAGAGTGTGAGTCTCCCTTTCGCCCTCAGCGAGCCAGCGTGGGACCACTC 593

DB 25 CCGGATGAAGAGTGTGAGTCTCCCTTTCGCCCTCAGCGAGCCAGCGTGGGACCACTC 84

```
QY 594 TTCCCGGAGCAGGCGCCCTGGGGGCACTTCTCAGGCAGACAGATTGATTTC 653
Db 85 TTCCCGGAGCAGGCGCCCTGGGGGCACTTCTCAGGCAGACAGATTGATTTC 144
QY 654 GTTCGGGATCTTGTCCTCCACCCAGTTTGTGTGCGCATCATCGGAAGAGGGCTTGAC 713
Db 145 GCTGCGGATCTTGTCCTCCACCCAGTTTGTGTGCGCATCATCGGAAGAGGGCTTGAC 204
QY 714 CATAAAGAACATCACTAAGCAGACCCAGTCGCCGGGTAGATATCCATAGAAAAGAGAACTC 773
Db 205 CATAAAGAACATCACTAAGCAGACCCAGTCGCCGGGTAGATATCCATAGAAAAGAGAACTC 264
QY 774 TGGAGTGCAGAGAACGCTGTACCAATCCATGCAACCCAGAGGGGACCTTCTGAAGCATG 833
Db 265 TGGAGTGCAGAGAACGCTGTACCAATCCATGCAACCCAGAGGGGACCTTCTGAAGCATG 324
QY 834 CGCGATGATCTTGAATCATGCAGAAAGAGGAGATGAGACCAACTAGCCGAGAGAT 893
Db 325 CGCGATGATCTTGAATCATGCAGAAAGAGGAGATGAGACCAACTAGCCGAGAGAT 384
QY 894 TCCTCTGAAAATCTTGGGCACAACTGCTTGGTGGGAAGCTGATTTGAAAAGAGGCGAG 953
Db 385 TCCTCTGAAAATCTTGGGCACAACTGCTTGGTGGGAAGCTGATTTGAAAAGAGGCGAG 444
QY 954 AAATTTGAAGAAATTTGAATCAATGAACAGGAGCAGATGATGATCAATCTCATCTTTGCAGGA 1013
Db 445 AAATTTGAAGAAATTTGAATCAATGAACAGGAGCAGATGATGATCAATCTCATCTTTGCAGGA 504
QY 1014 TTTGACATATACACCCGGAAGAACCATCTGTTGAAGGGCACAGTTTGAGGCCCTGTGC 1073
Db 505 TTTGACATATACACCCGGAAGAACCATCTGTTGAAGGGCACAGTTTGAGGCCCTGTGC 564
QY 1074 CAGTGTGATAGATAGATTATGAAGAA 1100
Db 565 CAGTGTGATAGATAGATTATGAAGAA 591

RESULT 15
AC126373
LOCUS
DEFINITION Homo sapiens chromosome 18 clone CTD-2307113 map 18, LOW-PASS
SEQUENCE SAMPLING.
AC126373
VERSION AC126373.1 GI:21699324
KEYWORDS HTG; HTGS_PHASE0.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 50320)
Birren,B., Nusbaum,C. and Lander,E.
Homo sapiens chromosome 18, clone CTD-2307113
Unpublished
REFERENCE
1 (bases 1 to 50320)
Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
Barns,N., Bastien,V., Bloom,T., Boguslavsky,L., Boukhvalter,B.,
Camara,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
Cook,A., Cooke,P., DeArlano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J.,
Gardyna,S., Gord,S., Graham,L., Grand-pierre,N., Hagos,B.,
Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A.,
Karacas,A., Kells,C., Lander,T., Levine,R., Lindblad-Toh,K.,
Liu,G., MacLean,C., Macdonald,P., Major,J., Matthews,C.,
McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,
O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,
Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P.,
Roman,J., Roy,A., Schauer,S., Schupbach,R., Seaman,S., Severy,P.,
Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J.,
Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,
Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
TITLE
```

JOURNAL  
COMMENT

Submitted (05-JUL-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
http://ftp.genome.washington.edu/BM/RepeatMasker.html  
-----  
Center: Whitehead Institute/ MIT Center for Genome Research  
Genome Center  
Center code: WIER  
Web site: http://www-seq.wi.mit.edu  
Contact: sequence.submissions@genome.wi.mit.edu  
-----  
Project Information  
Center project name: L27661  
Center clone name: 2307\_I\_13  
-----

\* NOTE: This record contains 59 individual  
\* sequencing reads that have not been assembled into  
\* contigs. Runs of N are used to separate the reads  
\* and the order in which they appear is completely  
\* arbitrary. Low-pass sequence sampling is useful for  
\* identifying clones that may be gene-rich and allows  
\* overlap relationships among clones to be deduced.  
\* However, it should not be assumed that this clone  
\* will be sequenced to completion. In the event that  
\* the record is updated, the accession number will  
\* be preserved.

1 766: contig of 766 bp in length  
\* 767 866: gap of 100 bp  
\* 867 1574: contig of 708 bp in length  
\* 1575 1674: gap of 100 bp  
\* 1675 2412: contig of 738 bp in length  
\* 2413 2512: gap of 100 bp  
\* 2513 3283: contig of 771 bp in length  
\* 3284 3383: gap of 100 bp  
\* 3384 4124: contig of 741 bp in length  
\* 4125 4224: gap of 100 bp  
\* 4225 4938: contig of 714 bp in length  
\* 4939 5038: gap of 100 bp  
\* 5039 5820: contig of 782 bp in length  
\* 5821 5920: gap of 100 bp  
\* 5921 6685: contig of 765 bp in length  
\* 6686 6785: gap of 100 bp  
\* 6786 7580: contig of 795 bp in length  
\* 7581 7680: gap of 100 bp  
\* 7681 8432: contig of 752 bp in length  
\* 8433 8532: gap of 100 bp  
\* 8533 9281: contig of 749 bp in length  
\* 9282 9381: gap of 100 bp  
\* 9382 10136: contig of 755 bp in length  
\* 10137 10236: gap of 100 bp  
\* 10237 10967: contig of 731 bp in length  
\* 10968 11067: gap of 100 bp  
\* 11068 11787: contig of 720 bp in length  
\* 11788 11887: gap of 100 bp  
\* 11888 12655: contig of 768 bp in length  
\* 12656 12755: gap of 100 bp  
\* 12756 13497: contig of 742 bp in length  
\* 13498 13597: gap of 100 bp  
\* 13598 14365: contig of 768 bp in length  
\* 14366 14665: gap of 100 bp  
\* 14666 15203: contig of 738 bp in length  
\* 15204 15303: gap of 100 bp  
\* 15304 16010: contig of 707 bp in length  
\* 16011 16110: gap of 100 bp  
\* 16111 16872: contig of 762 bp in length  
\* 16873 16972: gap of 100 bp  
\* 16973 17115: contig of 743 bp in length  
\* 17116 17815: gap of 100 bp  
\* 17816 18587: contig of 772 bp in length  
\* 18588 18687: gap of 100 bp  
\* 18688 19464: contig of 777 bp in length  
\* 19465 19564: gap of 100 bp  
\* 19565 20303: contig of 739 bp in length  
\* 20304 20403: gap of 100 bp

20404 21135: contig of 732 bp in length  
21235: gap of 100 bp  
21236 21999: contig of 764 bp in length  
21236 22099: gap of 100 bp  
22000 22846: contig of 747 bp in length  
22100 22946: gap of 100 bp  
22847 23709: contig of 763 bp in length  
22947 23809: gap of 100 bp  
23710 24525: contig of 716 bp in length  
23810 24626: gap of 100 bp  
24526 25371: contig of 746 bp in length  
24626 25471: gap of 100 bp  
25372 26239: contig of 768 bp in length  
25472 26339: gap of 100 bp  
26240 27090: contig of 751 bp in length  
26340 27190: gap of 100 bp  
27091 27946: contig of 756 bp in length  
27191 28046: gap of 100 bp  
27947 28808: contig of 762 bp in length  
28047 28908: gap of 100 bp  
28809 29669: contig of 761 bp in length  
29670 30550: contig of 781 bp in length  
29670 30550: gap of 100 bp  
30551 31423: contig of 773 bp in length  
30551 31423: gap of 100 bp  
31424 32279: contig of 756 bp in length  
31524 32379: gap of 100 bp  
32280 33129: contig of 750 bp in length  
32380 33229: gap of 100 bp  
33130 33984: contig of 755 bp in length  
33230 34084: gap of 100 bp  
33985 34828: contig of 744 bp in length  
34085 34928: gap of 100 bp  
34829 35665: contig of 737 bp in length  
34929 35765: gap of 100 bp  
35666 36535: contig of 770 bp in length  
35766 36635: gap of 100 bp  
36536 37412: contig of 777 bp in length  
36636 37512: gap of 100 bp  
37413 38284: contig of 772 bp in length  
37513 38384: gap of 100 bp  
38285 39173: contig of 789 bp in length  
39174 39273: gap of 100 bp  
39274 40051: contig of 778 bp in length  
40052 40151: gap of 100 bp  
40152 40900: contig of 743 bp in length  
40901 41000: gap of 100 bp  
41001 41778: contig of 778 bp in length  
41779 42645: contig of 767 bp in length  
42646 42745: gap of 100 bp  
42746 43554: contig of 709 bp in length  
43555 44314: gap of 100 bp  
44315 44414: contig of 760 bp in length  
44415 45158: contig of 744 bp in length  
45159 45258: gap of 100 bp  
45259 46040: contig of 782 bp in length  
46041 46140: gap of 100 bp  
46141 46894: contig of 754 bp in length  
46895 46994: gap of 100 bp  
46995 47754: contig of 760 bp in length  
47755 47854: gap of 100 bp  
47855 48596: contig of 742 bp in length  
48597 49446: gap of 100 bp  
49447 49546: contig of 750 bp in length  
49547 50320: contig of 774 bp in length.

FEATURES  
Source

ORIGIN  
Query Match  
Best Local Similarity 16.6%; Score 565; DB 2; Length 50320;  
Matches 615; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1917 ACACCTGACAGAAATGAGACCAACGCGAGCCAGATCGGAGCAACCAAGACCATC 1976  
DB 37566 ACACCTGACAGAAATGAGACCAACGCGAGCCAGATCGGAGCAACCAAGACCATC 37625  
QY 1977 TGAGGAATGAGAACTCTGCGAGGCGGCGAGGACTCTGCGAGGCGGCTTGAGAACCC 2036  
DB 37626 TGAGGAATGAGAACTCTGCGAGGCGGCGAGGACTCTGCGAGGCGGCTTGAGAACCC 37685  
QY 2037 GGGCGGAGGAGGGGGGGGAGGTCAGCCAGGTTTGCAGAACCAACCGAGCCCGCTCC 2096  
DB 37686 GGGCGGAGGAGGGGGGGGAGGTCAGCCAGGTTTGCAGAACCAACCGAGCCCGCTCC 37745  
QY 2097 CGCCCCCAGGGCTTCTGAGGCTTTCAGGCTTCCACATCCACTCGGATCTCTCC 2156  
DB 37746 CGCCCCCAGGGCTTCTGAGGCTTTCAGGCTTCCACATCCACTCGGATCTCTCC 37805  
QY 2157 TGAATCTCCAGCAGCTATCCCTTTTATGTTGAACTAATAGGTGAACGTTCAAGGCC 2216  
DB 37806 TGAATCTCCAGCAGCTATCCCTTTTATGTTGAACTAATAGGTGAACGTTCAAGGCC 37865  
QY 2217 AAGCAAAATGCACACCCCTTTTCTGTCGCAAACTGCTCTGTACATGTGTACATATTA 2276  
DB 37866 AAGCAAAATGCACACCCCTTTTCTGTCGCAAACTGCTCTGTACATGTGTACATATTA 37925  
QY 2277 GAAAGGGAAGATGTTAAGATATGTGGCTGTGGGTTACACAGGCTGCTGAGCGGTAAT 2336  
DB 37926 GAAAGGGAAGATGTTAAGATATGTGGCTGTGGGTTACACAGGCTGCTGAGCGGTAAT 37985  
QY 2337 ATATTTAGAAATAATATATCAATAACTCAACTAACTCCAAATTTTAAATCAATATTA 2396  
DB 37986 ATATTTAGAAATAATATATCAATAACTCAACTAACTCCAAATTTTAAATCAATATTA 38045  
QY 2397 TTTTCTTTCTTTTAAAGAGAAAGCAGGCTTTTCTAGACTTTTAAAGAAATAGCTTTG 2456  
DB 38046 TTTTCTTTCTTTTAAAGAGAAAGCAGGCTTTTCTAGACTTTTAAAGAAATAGCTTTG 38105  
QY 2457 GGAGGCTCAGGCTGTAGAGAGGAGCTTTTGAGGCCACCCGACAAAATTTCAACCCAGAGG 2516  
DB 38106 GGAGGCTCAGGCTGTAGAGAGGAGCTTTTGAGGCCACCCGACAAAATTTCAACCCAGAGG 38165  
QY 2517 AAATCTCGTGGGAGG 2532  
DB 38166 AAATCTCGTGGGAGG 38181

Search completed: July 21, 2004, 16:45:49  
Job time : 12978 secs

Location/Qualifiers  
1. .50320  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 21, 2004, 12:23:17 ; Search time 1236 Seconds  
(without alignments)  
11727.226 Million cell updates/sec

Title: US-09-270-437D-6  
Perfect score: 3412  
Sequence: 1 ggcagcggaggcggaggagga.....aaccttgaaatgtttattt 3412

Scoring table: OLIGO\_NUC  
Gapop\_60.0 , Gapext 60.0

Searched: 3373863 seqs, 2124099041 residues

Word size : 0

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N Geneseq\_29Jan04:\*  
1: Geneseqn1980s:\*  
2: Geneseqn1990s:\*  
3: Geneseqn2000s:\*  
4: Geneseqn2001as:\*  
5: Geneseqn2001bs:\*  
6: Geneseqn2002s:\*  
7: Geneseqn2003as:\*  
8: Geneseqn2003bs:\*  
9: Geneseqn2003cs:\*  
10: Geneseqn2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	3412	100.0	3412	3	Aaz36152 DNA encod
2	2953	86.5	3694	4	Aas26150 Human CDN
3	2953	86.5	3694	7	Abx73491 Human nov
4	2143	62.8	3283	3	Aaz36154 An altern
5	1977	57.9	2010	5	Aas70981 DNA encod
6	1810	53.0	3667	5	Aas70982 DNA encod
7	1810	53.0	3667	6	Abx76442 CDNA enco
8	1496	43.8	2290	4	Aas26153 Human CDN
9	1496	43.8	2290	7	Abx73494 Human nov
10	567	16.6	710	3	Aa02565 Human col
11	401	11.8	1186	4	Aas26570 Human CDN
12	401	11.8	1186	7	Abx73911 Human nov
13	389	11.4	822	4	Aas26151 Human CDN
14	389	11.4	822	7	Abx73492 Human nov
15	347	10.2	364	3	Aac03267 Human sec
16	290	8.5	424	3	Aah31011 Human col
17	285	8.4	300	3	Aa01526 Human col
18	264	7.7	264	4	Aas24943 Human ova
19	264	7.7	364	5	Aah83588 Human ova
20	263	7.7	263	4	Aas25126 Human ova
21	263	7.7	263	4	Aas25356 Human ova
22	263	7.7	378	5	Aah83772 Human ova
23	222	6.5	222	4	Aas58000 CDNA #676

24	222	6.5	222	4	AAS58609	CDNA #128
25	218	6.4	372	9	ADD71082	Human IGF
26	167	4.9	218	4	AAS58739	CDNA #141
27	156	4.6	306	2	AAV87489	EST clone
28	143	4.2	402	8	ACH18272	Human adu
29	141	4.1	184	2	Aah86031	Human sin
30	140	4.1	243	4	AAS57938	CDNA #614
31	135	4.0	1707	7	ACA90176	CDNA enco
32	129	3.8	228	3	AAA41968	Human sec
33	93	2.7	253	4	AAS58053	CDNA #729
34	92	2.7	598	5	ABV56560	Human pro
35	83	2.4	380	6	ABL67315	Thyroid c
36	83	2.4	380	6	ABN94637	Gene #113
37	68	2.0	412	6	ABL64217	Stomach c
38	62	1.8	145	4	AAS26568	Human CDN
39	62	1.8	145	7	ABX73909	Human nov
40	56	1.6	1078	3	AAA02581	Human col
41	55	1.6	564	8	AA157525	Human IGF
42	42	1.2	419	4	AA186673	Human pol
43	37	1.1	495	8	ACH39118	Human foe
44	27	0.8	1404	6	ABQ22104	Oligonuc1
45	27	0.8	1404	6	ABQ22105	Oligonuc1

ALIGNMENTS

RESULT 1  
AAZ36152  
ID AAZ36152 standard; DNA; 3412 BP.  
XX  
AC AAZ36152;  
XX  
DT 11-FEB-2000 (first entry)  
XX  
DE DNA encoding cancer associated antigen KOC-3.  
XX  
KW Cancer associated antigen; KOC-3; cancer; vaccine; CT7; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO9954738-A1.  
XX  
PD 28-OCT-1999.  
XX  
PF 16-MAR-1999; 99WO-US005766.  
XX  
PR 17-APR-1998; 98US-00061709.  
XX  
PA (LUDW-) LUDWIG INST CANCER RES.

Chen Y, Gure A, Tsang S, Stockert E, Jager E, Knuth A, Old LJ;  
WPI; 2000-013284/01.  
XX  
Nucleotides representing cancer-associated genes, used to develop products for the diagnosis, monitoring and treatment of cancers.  
Claim 55; Page 41; 44pp; English.  
XX  
The present sequence represents a cancer associated antigen gene designated KOC-3. The specification also describes a cancer associated antigen designated CT7. The CT7 polynucleotide was isolated from SK-MEL-37 melanoma cells. The polypeptide has some homology with MAGS-10, CC limited to about 210 carboxy terminal amino acids. The amino terminal of the protein has a repetitive pattern, with repeats rich in serine, CC proline, glutamine and leucine, and an almost invariable core of the CC peptide given in AAY43877. The CT7 polypeptide can be processed to CC peptides which provoke lysis by cytolytic T cells. The polynucleotides CC and polypeptides can be used for treating a cancerous condition and CC screening for or diagnosing cancerous conditions. The cancer associated CC antigens can be used as an immunogenic or vaccine composition with an CC adjuvant, e.g. a cytokine, a saponin, or granulocyte macrophage-colony

CC	stimulating factor (GM-CSF)									
XX										
SQ	Sequence 3412 BP; 970 A; 887 C; 804 G; 750 T; 0 U; 1 Other;									
	Query Match 100.0%; Score 3412; DB 3; Length 3412;									
	Best Local Similarity 100.0%; Pred. No. 0;									
	Matches 3412; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY	1	GGCAGCGAGAGGCGAGCGCGGTACCGGCCCGGGGAGCGCGGGCTCTCGGGG	60							
DB	1	GGCAGCGAGAGGCGAGAGCGCGGGTACCGGGCGGGGAGCGCGGGCTCTCGGGG	60							
QY	61	AAGAGACGGATGATGAACAAAGCTTTACATCGGAACTCTGAGCCCGCGTCAACGCCGAC	120							
DB	61	AAGAGACGGATGATGAACAAAGCTTTACATCGGAACTCTGAGCCCGCGTCAACGCCGAC	120							
QY	121	GACCTCCGGCAGCTTTTGGGACAGGAAGCTGCCCCCTGGCGGACAGTCTCTGTGAAG	180							
DB	121	GACCTCCGGCAGCTTTTGGGACAGGAAGCTGCCCCCTGGCGGACAGTCTCTGTGAAG	180							
QY	181	TCCGGCTACGCTTTCGTGGAATACCCCGACAGAACTGGGCCATCCGCGCCATCGAGACC	240							
DB	181	TCCGGCTACGCTTTCGTGGAATACCCCGACAGAACTGGGCCATCCGCGCCATCGAGACC	240							
QY	241	CTCTCGGGTAAAGTGGAAATGCATGGGAAATCATGGAAGTTGATCTCAGTCTCTAAA	300							
DB	241	CTCTCGGGTAAAGTGGAAATGCATGGGAAATCATGGAAGTTGATCTCAGTCTCTAAA	300							
QY	301	AAGCTAAGGACAGGAAATTCAGATTCGAAACATCCCTCTCAGCTCAGTGGAGGTG	360							
DB	301	AAGCTAAGGACAGGAAATTCAGATTCGAAACATCCCTCTCAGCTCAGTGGAGGTG	360							
QY	361	TTGGATGGACTTTTGGCTCAATATGGGACAGTGGAGAAATGGAACAACTCAACAGAC	420							
DB	361	TTGGATGGACTTTTGGCTCAATATGGGACAGTGGAGAAATGGAACAACTCAACAGAC	420							
QY	421	ACAGAAACCGCGTGTGCAACGTCAATATCAACAGAGAAAGCAAAATAGCCATG	480							
DB	421	ACAGAAACCGCGTGTGCAACGTCAATATCAACAGAGAAAGCAAAATAGCCATG	480							
QY	481	GAGAACTAAGCGGACATCAGTTTGAGAACTACCTCTCAAGATTCCTACATCCGGAT	540							
DB	481	GAGAACTAAGCGGACATCAGTTTGAGAACTACCTCTCAAGATTCCTACATCCGGAT	540							
QY	541	GAAGAGGTGAGCTCCCTTCCGCCCTCAGCGAGCCAGCGTGGGACCACTCTTCCCGG	600							
DB	541	GAAGAGGTGAGCTCCCTTCCGCCCTCAGCGAGCCAGCGTGGGACCACTCTTCCCGG	600							
QY	601	GAGCAAGGCCACGCCCTTGGGGACATTCCTCAGGCGCAGACAGATGATTTCCCGCTCGG	660							
DB	601	GAGCAAGGCCACGCCCTTGGGGACATTCCTCAGGCGCAGACAGATGATTTCCCGCTCGG	660							
QY	661	ATCCTGCTCCCAACCCAGTTCCTGCGGCTAGATATCCATGAAAGAGAACTCTGGAGCT	720							
DB	661	ATCCTGCTCCCAACCCAGTTCCTGCGGCTAGATATCCATGAAAGAGAACTCTGGAGCT	720							
QY	721	AACATCACTAAGCAGACCCAGTCCCGGTAGATATCCATGAAAGAGAACTCTGGAGCT	780							
DB	721	AACATCACTAAGCAGACCCAGTCCCGGTAGATATCCATGAAAGAGAACTCTGGAGCT	780							
QY	781	GCAGAGAGCTGTACCATTCATGCCCCAGAGGGACTTCTGAAGCATGCCGATG	840							
DB	781	GCAGAGAGCTGTACCATTCATGCCCCAGAGGGACTTCTGAAGCATGCCGATG	840							
QY	841	ATTCTTGAATCATGAGAAAGAGGAGGAGTGAACCAAACTAGCCGGAAGATTTCTCTG	900							
DB	841	ATTCTTGAATCATGAGAAAGAGGAGGAGTGAACCAAACTAGCCGGAAGATTTCTCTG	900							
QY	901	AAAATCTTGCAACAATGGCTTGGTTGGAAGTCTGATTTGAAAAGGCGAGAAATTTG	960							
DB	901	AAAATCTTGCAACAATGGCTTGGTTGGAAGTCTGATTTGAAAAGGCGAGAAATTTG	960							
QY	961	AAGAAATTTGAACATGAAACAGGACCAAGATTAACATCTCATCTTTGAGGATTTGAGC	1020							

DB	961	AAGAAATTTGAACATGAAACAGGACCAAGATAACAATCTCATCTTTGAGGATTTGAGC	1020							
QY	1021	ATATCAACCCGGAAAGAACCATCACTGTGAAGGCGACAGATTGAGCCCTGTGCCAGTGT	1080							
DB	1021	ATATCAACCCGGAAAGAACCATCACTGTGAAGGCGACAGTTGAGCCCTGTGCCAGTGT	1080							
QY	1081	GAGATAGAGATTATGAAGAAGCTGGTGAGGCTTTTGAANAATGATATGCTGCTGTTAAC	1140							
DB	1081	GAGATAGAGATTATGAAGAAGCTGGTGAGGCTTTTGAANAATGATATGCTGCTGTTAAC	1140							
QY	1141	CAACAAGCCAATCTGATCCCAAGGTTGAACCTTCAGCGCACCTGGCATCTTTTCAACAGA	1200							
DB	1141	CAACAAGCCAATCTGATCCCAAGGTTGAACCTTCAGCGCACCTGGCATCTTTTCAACAGA	1200							
QY	1201	CTGTCGCTGCTATCTCCACAGCAGGCGCCCGCGAGCTCCCGCGCTGCCCTTACCAC	1260							
DB	1201	CTGTCGCTGCTATCTCCACAGCAGGCGCCCGCGAGCTCCCGCGCTGCCCTTACCAC	1260							
QY	1261	CCCTTCATACCCACTCCGGATCTTCTCCAGCCTGTACCCCGCATCACAGTTTGGCCG	1320							
DB	1261	CCCTTCATACCCACTCCGGATCTTCTCCAGCCTGTACCCCGCATCACAGTTTGGCCG	1320							
QY	1321	TTCCCGCATCATCTCTTATCCAGAGCAGGAGATTGTGAATCTCTTCATCCCAACCCAG	1380							
DB	1321	TTCCCGCATCATCTCTTATCCAGAGCAGGAGATTGTGAATCTCTTCATCCCAACCCAG	1380							
QY	1381	GCTGTGGCGCATCATCTCGGAAAGAGGGGCGACACATCAACAGCTGGCGAGATTGCC	1440							
DB	1381	GCTGTGGCGCATCATCTCGGAAAGAGGGGCGACACATCAACAGCTGGCGAGATTGCC	1440							
QY	1441	GGAGCTCTATCAAGATTGCCCCCTGCGGAGGCCCGACGCTCAGCGAAAGATGGTCTC	1500							
DB	1441	GGAGCTCTATCAAGATTGCCCCCTGCGGAGGCCCGACGCTCAGCGAAAGATGGTCTC	1500							
QY	1501	ATCACCGGGCCACCGGAAAGCCAGTTCAGAGCCCGAGGACGATCTTTGGGAAACTGAAA	1560							
DB	1501	ATCACCGGGCCACCGGAAAGCCAGTTCAGAGCCCGAGGACGATCTTTGGGAAACTGAAA	1560							
QY	1561	GAGAAAACTTTCTTTAAACCCCAAGAGAGTGAAGCTGGAAGCGCATATCAGAGTGCC	1620							
DB	1561	GAGAAAACTTTCTTTAAACCCCAAGAGAGTGAAGCTGGAAGCGCATATCAGAGTGCC	1620							
QY	1621	TCCTTCCACAGCTGCGCGGTGATTGGCAAGGTGGCAAGCCGTAAGACGACTCAGAAC	1680							
DB	1621	TCCTTCCACAGCTGCGCGGTGATTGGCAAGGTGGCAAGCCGTAAGACGACTCAGAAC	1680							
QY	1681	TTAACCCAGTGCAGAACTCATCGTCCCTGACCAACCGCAGATGAAATGAGAAAGTG	1740							
DB	1681	TTAACCCAGTGCAGAACTCATCGTCCCTGACCAACCGCAGATGAAATGAGAAAGTG	1740							
QY	1741	ATCGTCAGAAATTTATCGGCACTTCTTTGCTAGCCAGACTGCAAGCGCAGATCAGGAA	1800							
DB	1741	ATCGTCAGAAATTTATCGGCACTTCTTTGCTAGCCAGACTGCAAGCGCAGATCAGGAA	1800							
QY	1801	ATTGTACAACAGAGTGAAGCAGGAGCAGAAATACCTCAGGAGTTCGCTCAGACGC	1860							
DB	1801	ATTGTACAACAGAGTGAAGCAGGAGCAGAAATACCTCAGGAGTTCGCTCAGACGC	1860							
QY	1861	AGCAAGTGAAGCTTCCACAGGCCACAGCAAAACCAACCGATGAATGTAGCCCTTCCAC	1920							
DB	1861	AGCAAGTGAAGCTTCCACAGGCCACAGCAAAACCAACCGATGAATGTAGCCCTTCCAC	1920							
QY	1921	CTCACGAATGAGCAACCAACCGCAGCCAGATCGGAGCAAAACCAACCATCTGAG	1980							
DB	1921	CTCACGAATGAGCAACCAACCGCAGCCAGATCGGAGCAAAACCAACCATCTGAG	1980							
QY	1981	GAATGAGAGTCTTCGCGAGCGCCAGGGACTCTGCGAGGCCCTGAGAAACCCAGGGGC	2040							
DB	1981	GAATGAGAGTCTTCGCGAGCGCCAGGGACTCTGCGAGGCCCTGAGAAACCCAGGGGC	2040							
QY	2041	CGAGAGGGCGGGGAGGTGAGCCAGGTTTGCAGAAACCAACCGAGCCCGCTCCCGCC	2100							

Db 2041 CGAGGAGGGGGGGAAGTGCACCGAGGTTTGGCCAGAACACCGAGCGCCGCTCCGGCC 2100  
Qy 2101 CCCAGGGGCTTCTGCGAGGCTTACGCCATCCACTTCCACCATCCACTCGGATCTCTCTGAA 2160  
Db 2101 CCCAGGGCTTCTGCGAGGCTTACGCCATCCACTTCCACCATCCACTCGGATCTCTCTGAA 2160  
Qy 2161 CTCGCCAGCGCTATCCCTTTTGTAGTTGAACATAACATAGGTGAACGTTTCAAGCCCAAGC 2220  
Db 2161 CTCGCCAGCGCTATCCCTTTTGTAGTTGAACATAACATAGGTGAACGTTTCAAGCCCAAGC 2220  
Qy 2221 AAAATGCACACCCCTTTTCTGTGCGCAATCGTCTCTGTATCATCTGTGTACATATTAGAAA 2280  
Db 2221 AAAATGCACACCCCTTTTCTGTGCGCAATCGTCTCTGTATCATCTGTGTACATATTAGAAA 2280  
Qy 2281 GGAAGATGTTAAGATATGTGGCTGTGGGTTACACAGGCTGCTGCAGCGGTAATATAT 2340  
Db 2281 GGAAGATGTTAAGATATGTGGCTGTGGGTTACACAGGCTGCTGCAGCGGTAATATAT 2340  
Qy 2341 TTTAGAAATATATATCAATACTCAACTCACTCACTCACTCACTCACTCACTCACTCACTCACT 2400  
Db 2341 TTTAGAAATATATATCAATACTCAACTCACTCACTCACTCACTCACTCACTCACTCACTCACT 2400  
Qy 2401 TTTTCTTTTAAAGAGAAAGCAGGCTTTTCTAGACTTTTAAAGATAAAGTCTTTGGGAG 2460  
Db 2401 TTTTCTTTTAAAGAGAAAGCAGGCTTTTCTAGACTTTTAAAGATAAAGTCTTTGGGAG 2460  
Qy 2461 GTCTACGGTGTAGAGGAGCTTTTGGGCCACCGGCACCAAAATTTCCACAGAGGGAAAT 2520  
Db 2461 GTCTACGGTGTAGAGGAGCTTTTGGGCCACCGGCACCAAAATTTCCACAGAGGGAAAT 2520  
Qy 2521 CTGCTCGAAGGACTCAACGAGCTTCTGATCAGCTGTGTATGTATGCAACAGAGGATA 2580  
Db 2521 CTGCTCGAAGGACTCAACGAGCTTCTGATCAGCTGTGTATGTATGCAACAGAGGATA 2580  
Qy 2581 CGCTCTCTTGAAGAGAAACTCTGCACCTCTCATGCCCTGTAGCTCTATACACCCATT 2640  
Db 2581 CGCTCTCTTGAAGAGAAACTCTGCACCTCTCATGCCCTGTAGCTCTATACACCCATT 2640  
Qy 2641 TCTCTTGTCTCAGGTTTTAACTGGTTTTTGTGCATCTGCTATATATTTCTGTCT 2700  
Db 2641 TCTCTTGTCTCAGGTTTTAACTGGTTTTTGTGCATCTGCTATATATTTCTGTCT 2700  
Qy 2701 CTCTCTGTTATCTCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 2760  
Db 2701 CTCTCTGTTATCTCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 2760  
Qy 2761 TTTCTCTATCCCTCCATCTCAATCCCGTATCTAGCCACCCGCCCCGCCAGGCAAGCA 2820  
Db 2761 TTTCTCTATCCCTCCATCTCAATCCCGTATCTAGCCACCCGCCCCGCCAGGCAAGCA 2820  
Qy 2821 GTGCTCTGAGTATCACATCAACAAAAAGGAAACAAAAAGGAAACACAAAAACCAAGCCCTCAA 2880  
Db 2821 GTGCTCTGAGTATCACATCAACAAAAAGGAAACAAAAAGGAAACACAAAAACCAAGCCCTCAA 2880  
Qy 2881 CTTCACCTTGGTTACTCAAGAGACAGTCAATGCTACTGTCTAGCGTTTGGAG 2940  
Db 2881 CTTCACCTTGGTTACTCAAGAGACAGTCAATGCTACTGTCTAGCGTTTGGAG 2940  
Qy 2941 AGGAAAAAGGAAACCCACCAACCAACCAATCAACCAACCAACCAACCAACCAACCAACCAATG 3000  
Db 2941 AGGAAAAAGGAAACCCACCAACCAACCAATCAACCAACCAACCAACCAACCAACCAACCAATG 3000  
Qy 3001 AAAGAAATGATTTTGTCTTTTGTGATTTGATTAAGCCATCAATATTACAGAAATG 3060  
Db 3001 AAAGAAATGATTTTGTCTTTTGTGATTTGATTAAGCCATCAATATTACAGAAATG 3060  
Qy 3061 ATTCTCTTTTAAAAAAAATGTGGAGAAATAGAAATTTACAGGTTTGTGGCC 3120  
Db 3061 ATTCTCTTTTAAAAAAAATGTGGAGAAATAGAAATTTACAGGTTTGTGGCC 3120  
Qy 3121 CAGGCGTTAAATTCACAGATTTTTTTTAAACAGAAAAACACACAGAGAAAGCTACCTCAG 3180  
Db 3121 CAGGCGTTAAATTCACAGATTTTTTTTAAACAGAAAAACACACAGAGAAAGCTACCTCAG 3180

## RESULT 2

AAS26150

ID AAS26150 standard; cDNA; 3694 BP.

XX AAS26150;

XX AC AC

XX DT 07-NOV-2001 (first entry)

XX Human cDNA encoding a novel secreted protein, Seq ID 329.

XX Human; immunosuppressive; antiarthritic; ss; antirheumatic; cytostatic;  
cardiant; vasotropic; cerebroprotective; nootropic; neuroprotective;  
antibacterial; vitucide; fungicide; ophthalmological; vulnery;  
secreted protein; rheumatoid arthritis; hyperproliferative disorder;  
cardiovascular disorder; cardiac arrest; cerebrovascular disorder;  
cerebral ischaemia; angiogenesis; nervous system disorder;  
Alzheimer's disease; infection; ocular disorder; corneal infection;  
wound healing; epithelial cell proliferation; skin ageing; food additive;  
preservative; antiproliferative.

Homo sapiens.

XX OS

XX XX

XX WO200155322-A2.

XX PD 02-AUG-2001.

XX PF 17-JAN-2001; 2001WO-US001341.

XX PR 31-JAN-2000; 2000US-0179065P.

XX PR 04-FEB-2000; 2000US-0180628P.

XX PR 24-FEB-2000; 2000US-0184664P.

XX PR 02-MAR-2000; 2000US-0186350P.

XX PR 16-MAR-2000; 2000US-0189874P.

XX PR 17-MAR-2000; 2000US-0190076P.

XX PR 18-APR-2000; 2000US-0198123P.

XX PR 19-MAY-2000; 2000US-0205515P.

XX PR 07-JUN-2000; 2000US-0209467P.

XX PR 28-JUN-2000; 2000US-0214886P.

XX PR 30-JUN-2000; 2000US-0215135P.

XX PR 07-JUL-2000; 2000US-0216647P.

XX PR 11-JUL-2000; 2000US-0217487P.

XX PR 11-JUL-2000; 2000US-0217496P.

XX PR 14-JUL-2000; 2000US-0218290P.

XX PR 26-JUL-2000; 2000US-0220963P.

XX PR 26-JUL-2000; 2000US-0220964P.

XX PR 14-AUG-2000; 2000US-0224518P.

XX PR 14-AUG-2000; 2000US-0224519P.

XX PR 14-AUG-2000; 2000US-0225213P.

XX PR 14-AUG-2000; 2000US-0225214P.

XX PR 14-AUG-2000; 2000US-0225266P.

XX PR 14-AUG-2000; 2000US-0225267P.

XX PR 14-AUG-2000; 2000US-0225268P.

XX PR 14-AUG-2000; 2000US-0225270P.

XX PR 14-AUG-2000; 2000US-0225447P.

PR 14-AUG-2000; 2000US-0225757P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0226688P.  
PR 22-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 12-SEP-2000; 2000US-0232081P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 25-SEP-2000; 2000US-0234998P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235836P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239935P.  
PR 13-OCT-2000; 2000US-0239937P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241826P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0246474P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.

PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249264P.  
PR 17-NOV-2000; 2000US-0249265P.  
PR 17-NOV-2000; 2000US-0249297P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 17-NOV-2000; 2000US-0249300P.  
PR 01-DEC-2000; 2000US-0250160P.  
PR 05-DEC-2000; 2000US-0251030P.  
PR 05-DEC-2000; 2000US-0251088P.  
PR 05-DEC-2000; 2000US-0256719P.  
PR 06-DEC-2000; 2000US-0251479P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.  
PR 08-DEC-2000; 2000US-0251989P.  
PR 11-DEC-2000; 2000US-0251990P.  
PR 11-DEC-2000; 2000US-0254097P.  
PR 05-JAN-2001; 2001US-0259678P.  
XX

(HUMA-) HUMAN GENOME SCI INC.

PI Rosen CA, Barash SC, Ruben SM;

XX WPI; 2001-488783/53.

XX P-PSDB; AAU16163.

PT New nucleic acid molecules encoding 461 human secreted proteins for  
PT diagnosing, preventing, treating or ameliorating medical conditions and  
PT used as food additives or preservatives.

XX Claim 1; SEQ ID NO 329; 980pp; English.

XX The invention relates to isolated nucleic acid molecules and their  
CC encoded secreted proteins. The nucleic acids and proteins are used to  
CC prevent, treat or ameliorate a medical condition in e.g. humans, mice,  
CC rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used  
CC in diagnosing a pathological condition or susceptibility to a  
CC pathological condition. Antibodies to the proteins can also be used in  
CC alleviating symptoms associated with the disorders and in diagnostic  
CC immunoassays e.g. radioimmunoassays or enzyme linked immunosorbent assays  
CC (ELISA). Disorders which are diagnosed or treated include autoimmune  
CC diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g.  
CC neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac  
CC arrest, cerebrovascular disorders e.g. cerebral ischaemia, angiogenesis,  
CC nervous system disorders e.g. Alzheimer's disease, infections caused by  
CC bacteria, viruses and fungi and ocular disorders e.g. corneal infection,  
CC and many other disorders listed in the specification. The polypeptides  
CC can also be used to aid wound healing and epithelial cell proliferation,  
CC to prevent skin aging due to sunburn, to maintain organs before  
CC transplantation, for supporting cell culture of primary tissues, to  
CC regenerate tissues and in chemotaxis. The polypeptides can also be used  
CC as a food additive or preservative to increase or decrease storage  
CC capabilities, fat content, lipid, protein, carbohydrate, vitamins,  
CC minerals, cofactors and other nutritional components. The present  
CC sequence encodes a novel secreted protein of the invention. Note: The

CC sequence data for this patent did not form part of the printed

Query Match		86.5%;	Score 2953;	DB 4;	Length 3694;
Best Local Similarity		99.9%;	Pred. No. 0;		
Matches 3303;		Conservative	0;	Mismatches	1;
				Indels	3;
				Gaps	2;
Qy	68	GGATGATGACAGCTTTACATCGGGAACCTGAGCCCGCGCTGACCGCGAGACCTCC	127		
Db	91	GGATGATGACAGCTTTACATCGGGAACCTGAGCCCGCGCTGACCGCGAGACCTCC	150		
Qy	128	GGCAGCTCTTTGGGGACAGGAAGCTGCCCTCGCGGACAGGTCCTCTCAAGTCCGGCT	187		
Db	151	GGCAGCTCTTTGGGGACAGGAAGCTGCCCTCGCGGACAGGTCCTCTCAAGTCCGGCT	210		
Qy	188	ACGCCCTTGGGACTACCCCGACAGAACTCGGCCATCCCGGCCATCGAGACCTCTCGG	247		
Db	211	ACGCCCTTGGGACTACCCCGACAGAACTCGGCCATCCCGGCCATCGAGACCTCTCGG	270		
Qy	248	GTAAGCTGGATTCATGCGGAATCATGGAAGTTGATTAATCAGTCTCTAAAAGCTAA	307		
Db	271	GTAAGCTGGATTCATGCGGAATCATGGAAGTTGATTAATCAGTCTCTAAAAGCTAA	330		
Qy	308	GGAGCAGGAAATTCAGATTGGAATTCGAAATCCCTCTCACCTGCGAGTGGAGGTGTGGATG	367		
Db	331	GGAGCAGGAAATTCAGATTGGAATTCGAAATCCCTCTCACCTGCGAGTGGAGGTGTGGATG	390		
Qy	368	GACTTTGGTCAATATGGGACAGTGGGAATGGAACCAAGTCAACACACACAGAA	427		
Db	391	GACTTTGGTCAATATGGGACAGTGGGAATGGAACCAAGTCAACACACACAGAA	450		
Qy	428	CCGCCGTGTCACAGTCACATATGCAACAGAGAGAGCAAAATAGCCATGGAGAGC	487		
Db	451	CCGCCGTGTCACAGTCACATATGCAACAGAGAGAGCAAAATAGCCATGGAGAGC	510		
Qy	488	TAAGCGGCACTCAGTTTGAAATCTACCTCTCAAGATTTCTACATCCCGATGAAAGG	547		
Db	511	TAAGCGGCACTCAGTTTGAAATCTACCTCTCAAGATTTCTACATCCCGATGAAAGG	570		
Qy	548	TGAGTCCCTTCGCCCTTCAGGAGCCAGGTCGGGACCACTCTTCCCGGAGCAAG	607		
Db	571	TGAGTCCCTTCGCCCTTCAGGAGCCAGGTCGGGACCACTCTTCCCGGAGCAAG	630		
Qy	608	GCCACGCCCTGGGGGCACTCTCAGGCGACAGATTTGATTTCCCGCTCGGATCCTGG	667		
Db	631	GCCACGCCCTGGGGGCACTCTCAGGCGACAGATTTGATTTCCCGCTCGGATCCTGG	690		
Qy	668	TCCCCACCCAGTTTGTGGTGCCATCATCGAAAGAGGGCTTGACCATAAAGAACATCA	727		
Db	691	TCCCCACCCAGTTTGTGGTGCCATCATCGAAAGAGGGCTTGACCATAAAGAACATCA	750		
Qy	728	CTAAGCAGACCCAGTCCCGGTAGATATCCATAGAAAAGAGAACTCTGGAGCTGCAGAG	787		
Db	751	CTAAGCAGACCCAGTCCCGGTAGATATCCATAGAAAAGAGAACTCTGGAGCTGCAGAG	810		
Qy	788	AGCTGTCCACATCCATGCCCCAGAGGGGACTCTGAGCATGCGCATGATCTTCTG	847		
Db	811	AGCTGTCCACATCCATGCCCCAGAGGGGACTCTGAGCATGCGCATGATCTTCTG	870		
Qy	848	AAATCATGCAAGAGAGGAGATGAGACCAAACTAGCCGAGAGATTCCTCTGAAAATCT	907		
Db	871	AAATCATGCAAGAGAGGAGATGAGACCAAACTAGCCGAGAGATTCCTCTGAAAATCT	930		
Qy	908	TGGCACAATATGGCTTGGGAAGACTGATTTGAAAAGAGGAGGAGAAATTTGAAGAAA	967		
Db	931	TGGCACAATATGGCTTGGGAAGACTGATTTGAAAAGAGGAGGAGAAATTTGAAGAAA	990		
Qy	968	TTGACATGAAACAGGAGCAAGATATCAATCTCTTTCAGGATTTGAGCATATACA	1027		
Db	991	TTGACATGAAACAGGAGCAAGATATCAATCTCTTTCAGGATTTGAGCATATACA	1050		
Qy	1028	ACCCGGAAGAACCATCACTGTGAAGGCGACAGTTGAGGCTGTGCCAGTCTGAGATAG	1087		
Db	1051	ACCCGGAAGAACCATCACTGTGAAGGCGACAGTTGAGGCTGTGCCAGTCTGAGATAG	1110		

Qy	1088	AGATTATGAAGAGCTGCGTGAGGCTTTGAAAATGATATGCTGGCTGTGTTAAACAAG	1147
Db	1111	AGATTATGAAGAGCTGCGTGAGGCTTTGAAAATGATATGCTGGCTGTGTTAAACAAG	1170
Qy	1148	CCAACTGTATCCAGAGGTTGAACCTCAGCGCACTTGGCATCTTTTCAACAGGAGTGTCCG	1207
Db	1171	CCAACTGTATCCAGAGGTTGAACCTCAGCGCACTTGGCATCTTTTCAACAGGAGTGTCCG	1230
Qy	1208	TGCTATCTCCACAGAGGCGCCCGGAGCTCCCCCGCTGCCCTTACACCCCTTCA	1267
Db	1231	TGCTATCTCCACAGAGGCGCCCGGAGCTCCCCCGCTGCCCTTACACCCCTTCA	1290
Qy	1268	CTAACCACTCCGATATCTTCTCAGGCTGTATCCCCCATCAACAGTTTGGCCCGTTCGCG	1327
Db	1291	CTAACCACTCCGATATCTTCTCAGGCTGTATCCCCCATCAACAGTTTGGCCCGTTCGCG	1350
Qy	1328	ATCATCACTCTTATCCAGAGCAGGAGATTGTGAATCTCTTCATCCCAACCCAGGCTGTGG	1387
Db	1351	ATCATCACTCTTATCCAGAGCAGGAGATTGTGAATCTCTTCATCCCAACCCAGGCTGTGG	1410
Qy	1388	GCGCCATCATCGGGAAGAGGGGCGCACATCAAAAGAGCTGGCGAGATTCCCGGAGCCT	1447
Db	1411	GCGCCATCATCGGGAAGAGGGGCGCACATCAAAAGAGCTGGCGAGATTCCCGGAGCCT	1470
Qy	1448	CTATCAAGATTGCCCCCTGCGGAAGGCCAGACGTCAAGCGAAAGATGGTCAATCACCG	1507
Db	1471	CTATCAAGATTGCCCCCTGCGGAAGGCCAGACGTCAAGCGAAAGATGGTCAATCACCG	1530
Qy	1508	GGCCACCGGAAGCCAGTTCAAGGCCCGAGGAGCGGATCTTTGGGAACCTGAAAGAGAA	1567
Db	1531	GGCCACCGGAAGCCAGTTCAAGGCCCGAGGAGCGGATCTTTGGGAACCTGAAAGAGAA	1590
Qy	1568	ACTTCTTTTAAACCCCAAGAGAGTGAAGCTGGAAGCGCATATCAGAGTCCCTCTTTCCA	1627
Db	1591	ACTTCTTTTAAACCCCAAGAGAGTGAAGCTGGAAGCGCATATCAGAGTCCCTCTTTCCA	1650
Qy	1628	CAGCTGCGCGGTGATTTGGCAAGAGTGGGCAAGACCGGTGAACGAACTGAACCA	1687
Db	1651	CAGCTGCGCGGTGATTTGGCAAGAGTGGGCAAGACCGGTGAACGAACTGAACCA	1710
Qy	1688	GTCCAGAGTCATCTGCTCGCTGACCAAAAGCGGAGTGAAGTGAAGTGAAGTGAAGTGA	1747
Db	1711	GTCCAGAGTCATCTGCTCGCTGACCAAAAGCGGAGTGAAGTGAAGTGAAGTGAAGTGA	1770
Qy	1748	GAATTATCGGGCACTTCTTTGCTAGCCAGACTGCAAGCGCAAGATGAGGAAATGTATC	1807
Db	1771	GAATTATCGGGCACTTCTTTGCTAGCCAGACTGCAAGCGCAAGATGAGGAAATGTATC	1830
Qy	1808	AAAGGTGAAGCGAGGAGCAGAAATACCTCAGGAGTGCCTTCAAGCGCAGCAAGT	1867
Db	1831	AAAGGTGAAGCGAGGAGCAGAAATACCTCAGGAGTGCCTTCAAGCGCAGCAAGT	1890
Qy	1868	GAGCTCCACAGCAGCCAGCAAAACAAAGGATGAATGTAGCCCTTCCACACCTGACAG	1927
Db	1891	GAGCTCCACAGCAGCCAGCAAAACAAAGGATGAATGTAGCCCTTCCACACCTGACAG	1950
Qy	1928	AATGAGACCAAAACCGAGCCAGATCGGAGCAAAACCAAGACCATCTGAGGAATGAG	1987
Db	1951	AATGAGACCAAAACCGAGCCAGATCGGAGCAAAACCAAGACCATCTGAGGAATGAG	2010
Qy	1988	AAGTCTCGGAGCGGCGAGGAGCTCTGCGGAGGCTGTGAGAACCCAGGGGCGGAGGAG	2047
Db	2011	AAGTCTCGGAGCGGCGAGGAGCTCTGCGGAGGCTGTGAGAACCCAGGGGCGGAGGAG	2070
Qy	2048	GGCGGGGAGGTCAGCCAGTTTGGCAGAACCAACCGAGCCCGCTTCCCGCCCCCAGG	2107
Db	2071	GGCGGGGAGGTCAGCCAGTTTGGCAGAACCAACCGAGCCCGCTTCCCGCCCCCAGG	2130
Qy	2108	GCTTCTGACGGCTTCAAGCCATCCACTTCCACTCCACTCCACTCCACTCCACTCCACT	2167
Db	2131	GCTTCTGACGGCTTCAAGCCATCCACTTCCACTCCACTCCACTCCACTCCACTCCACT	2190

QY 2168 GACGCTATCCCTTTTGTAGTGAACATACTAGTGAAAGCGTTTCAAGGCCAAGCAAAATGC 2227  
Db 2191 GACGCTATCCCTTTTGTAGTGAACATACTAGTGAAAGCGTTTCAAGGCCAAGCAAAATGC 2250  
QY 2228 ACACCCCTTTTCTGTGGCAAAATCGTCTCTGTACATGTGTGTACATATTAGAAAGGGAAGA 2287  
Db 2251 ACACCCCTTTTCTGTGGCAAAATCGTCTCTGTACATGTGTGTACATATTAGAAAGGGAAGA 2310  
QY 2288 TGTTAAGATATGTGGCTGTGGGTTACACAGGGTGCCCTGCAGCGGTAAATATATTATTAGAA 2347  
Db 2311 TGTTAAGATATGTGGCTGTGGGTTACACAGGGTGCCCTGCAGCGGTAAATATATTATTAGAA 2370  
QY 2348 ATAAATATCAATAAATACTCAACTCAACTCAATTTTAAATCAATTAATTAATTTTCTCT 2407  
Db 2371 ATAAATATCAATAAATACTCAACTCAACTCAATTTTAAATCAATTAATTAATTTTCTCT 2430  
QY 2408 TTTTAAAGAGAAGCAGGCTTTTCTAGACTTTTAAAGAAATTAAGTCTTTTGGAGGCTCTAC 2467  
Db 2431 TTTTAAAGAGAAGCAGGCTTTTCTAGACTTTTAAAGAAATTAAGTCTTTTGGAGGCTCTAC 2490  
QY 2468 GGTGTAGAGAGAGCTTTGAGGCCACCGCCACAAAATTCACCCAGAGGGAATCTCGTCG 2527  
Db 2491 GGTGTAGAGAGAGCTTTGAGGCCACCGCCACAAAATTCACCCAGAGGGAATCTCGTCG 2550  
QY 2528 GAAGGACACTCAGCGAGTCTGTGATCACTGTGTATGTCAACAGAAGGGATACCGTCTC 2587  
Db 2551 GAAGGACACTCAGCGAGTCTGTGATCACTGTGTATGTCAACAGAAGGGATACCGTCTC 2610  
QY 2588 CTTGAAGAGGAACCTCTGCTCACTCTCATGCTGTCTAGCTCATACACCAATTTCTCTTT 2647  
Db 2611 CTTGAAGAGGAACCTCTGCTCACTCTCATGCTGTCTAGCTCATACACCAATTTCTCTTT 2670  
QY 2648 GCTTCACAGGTTTTAAACTGGTTTTTGCATCTGCTATATAATTCCTGTCTCTCTCTG 2707  
Db 2671 GCTTCACAGGTTTTAAACTGGTTTTTGCATCTGCTATATAATTCCTGTCTCTCTCTG 2730  
QY 2708 TTTATCTCTCCCTCCCTCCCTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2767  
Db 2731 TTTATCTCTCCCTCCCTCCCTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2790  
QY 2768 ATCCCTCCATCTCAATCCGATCTACGCA--CCCCCGCCCCCGCCAGGCAAGCTGCT 2825  
Db 2791 ATCCCTCCATCTCAATCCGATCTACGCA--CCCCCGCCCCCGCCAGGCAAGCTGCT 2850  
QY 2826 CTGAGTATCATCATCACAAAGGAACAAAGCGAAACACAAACCAAGCTCTCACTTAC 2885  
Db 2851 CTGAGTATCATCATCACAAAGGAACAAAGCGAAACACAAACCAAGCTCTCACTTAC 2910  
QY 2886 ACTTGGTTACTCAAAAGAACAAAGTCAATGGTACTTGTCTAGCGTTTTGGAGAGAA 2945  
Db 2911 ACTTGGTTACTCAAAAGAACAAAGTCAATGGTACTTGTCTAGCGTTTTGGAGAGAA 2970  
QY 2946 AACAGGAACCCACAAACCAACCAATCAACCAAGAAATAATCCCAATCAAGA 3005  
Db 2971 AACAGGAACCCACAAACCAACCAATCAACCAAGAAATAATCCCAATCAAGA 3030  
QY 3006 ATGTAATTTGTCTTTTGCATTTTGGTGTATAAGCCATCAATATTCAGCAAAATGATTC 3065  
Db 3031 ATGTAATTTGTCTTTTGCATTTTGGTGTATAAGCCATCAATATTCAGCAAAATGATTC 3090  
QY 3066 TTTCTTTT-AAAAAATAATGTGGAGAAAGTAGAAATTTACCAAGGTTTGTGGCCAGG 3124  
Db 3091 TTTCTTTTAAAAAATAATGTGGAGAAAGTAGAAATTTTACCAAGGTTTGTGGCCAGG 3150  
QY 3125 GCGTTAAATTCACAGATTTTTTAAACGAGAAACACACAGAAAGCTACTCAGGTGT 3184  
Db 3151 GCGTTAAATTCACAGATTTTTTAAACGAGAAACACACAGAAAGCTACTCAGGTGT 3210  
QY 3185 TTTTACCTCAGCACCTCTCTGTGTGTTCCTTAGAGATTTTGTAAAGCTGATAGTTGG 3244  
Db 3211 TTTTACCTCAGCACCTCTCTGTGTGTTCCTTAGAGATTTTGTAAAGCTGATAGTTGG 3270  
QY 3245 AGCATTTTTTTTATTAATAAATAAGTGTGGAAAAAATAAGATATCAACTGCCA 3304

Db 3271 AGCATTTTTTTTATTAATAAATAAGTGTGGAAAAAATAAGATATCAACTGCCA 3330  
QY 3305 GCCTGGAGAGGTGACAGTCCAAAGTGTGCAACAGCTGTTCTGAATTGTTCTCCGCTAGCC 3364  
Db 3331 GCCTGGAGAGGTGACAGTCCAAAGTGTGCAACAGCTGTTCTGAATTGTTCTCCGCTAGCC 3390  
QY 3365 AAGAAC 3371  
Db 3391 AAGAAC 3397  
RESULT 3  
ABX73491  
ID ABX73491 standard; DNA; 3694 BP.  
XX ABX73491;  
AC ABX73491;  
DT 18-MAR-2003 (first entry)  
XX Human novel polynucleotide #319.  
DE Human; gene; ds; neural disorder; immune system disorder; renal disorder;  
KW muscular disorder; respiratory disease; reproductive disorder;  
KW gastrointestinal disorder; pulmonary disorder; cardiovascular disorder;  
KW hyperproliferative disorder; inflammatory disease; allergic reaction;  
KW blood related disorder; cancer; immunosuppressive; antiinflammatory;  
KW cardiovascular; nephrotropic; cytostatic; antiallergic; thrombolytic;  
KW haemostatic; antiarteriosclerotic.  
XX Homo sapiens.  
XX US2002132753-A1.  
XX 19-SEP-2002.  
XX 17-JAN-2001; 2001US-00764864.  
PR 31-JAN-2000; 2000US-0179065P.  
PR 04-FEB-2000; 2000US-0180628P.  
PR 28-JUN-2000; 2000US-0214886P.  
PR 07-JUL-2000; 2000US-0216647P.  
PR 07-JUL-2000; 2000US-0216880P.  
PR 11-JUL-2000; 2000US-0217487P.  
PR 11-JUL-2000; 2000US-0217496P.  
PR 14-JUL-2000; 2000US-0218290P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 26-JUL-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225447P.  
PR 14-AUG-2000; 2000US-0225757P.  
PR 22-AUG-2000; 2000US-0226868P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.

PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239335P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 08-DEC-2000; 2000US-0251866P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.  
XX  
PA (ROSE/) ROSEN C A.  
PA (RUBE/) RUBEN S M.  
PA (BARA/) BARASH S C.  
XX  
PI Rosen CA, Ruben SM, Barash SC;  
DR WPI; 2003-147444/14.  
DR P-PSDB; ABUS5231.  
XX  
XX  
PT New polypeptides and nucleic acids, useful in gene therapy for treating,  
PT inhibiting or preventing e.g. neural, immune system, muscular,  
PT respiratory, reproductive, gastrointestinal, pulmonary, cardiovascular or  
PT renal disorders.  
XX  
PS Claim 1; SEQ ID NO 329; 402pp; English.  
XX  
CC The invention relates to human novel polypeptides and their associated  
CC polynucleotides. The polypeptides and polynucleotides are useful in gene  
CC therapy for treating, inhibiting or preventing neural disorders, immune  
CC system disorders (e.g. systemic lupus erythematosus, rheumatoid arthritis  
CC and multiple sclerosis), muscular disorders, respiratory diseases (e.g.  
CC nasal vestibulitis, nasal polyps and sinusitis), reproductive disorders,  
CC gastrointestinal disorders, pulmonary disorders, cardiovascular disorders  
CC (e.g. congenital heart defects, Ebstein's anomaly and hypoplastic left  
CC heart syndrome), renal disorders (e.g. acute kidney failure and end-stage  
CC renal disease), hyperproliferative disorders (e.g. Hodgkin's disease and  
CC leukemia), inflammatory diseases (e.g. septic shock, bursitis and  
CC appendicitis), allergic reactions and conditions (e.g. asthma), blood  
CC related disorders (e.g. thrombosis, atherosclerosis and myocardial  
CC infarction) and cancerous diseases. Sequences ABX73173-ABX74167 represent  
CC human novel polynucleotides of the invention  
XX  
SQ Sequence 3694 BP; 1103 A; 934 C; 839 G; 817 T; 0 U; 1 Other;

Query Match 86.5%; Score 2953; DB 7; Length 3694;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 3303; Conservative 0; Mismatches 1; Indels 3; Gaps 2;

QY 68 GGATGATGAACAGCTTTACATCGGGAACCTGAGCCCGCGGTACCGCGGACAGCTCC 127  
DB 91 GGATGATGAACAGCTTTACATCGGGAACCTGAGCCCGCGGTACCGCGGACAGCTCC 150  
QY 128 GGCAGCTCTTTGGGGACAGGAAGCTGCCCTTGGCGGACAGGTCCTGCTGAAGTCCGGCT 187  
DB 151 GGCAGCTCTTTGGGGACAGGAAGCTGCCCTTGGCGGACAGGTCCTGCTGAAGTCCGGCT 210  
QY 188 ACGCTTCGTGGACTACCCGACAGAACTGGGCGCATCCGGGCCATCGAGACCTCTTCGG 247  
DB 211 ACGCTTCGTGGACTACCCGACAGAACTGGGCGCATCCGGGCCATCGAGACCTCTTCGG 270  
QY 248 GTAAGTGGAAATTCATGGGAATCATGGAGTTGATCTCAGTCTCTAAAAGCTAA 307  
DB 271 GTAAGTGGAAATTCATGGGAATCATGGAGTTGATCTCAGTCTCTAAAAGCTAA 330  
QY 308 GGACAGGAAATTCAGATTCGAACATCCCTCTCACCTGCAGTGGGAGGTGTGGATG 367  
DB 331 GGACAGGAAATTCAGATTCGAACATCCCTCTCACCTGCAGTGGGAGGTGTGGATG 390

QY 368 GACTTTTGGCTCAATATGGGACAGTGGAGAAATGTGAAACAAGTCAACACAGACACAGAA 427  
DB 391 GACTTTTGGCTCAATATGGGACAGTGGAGAAATGTGAAACAAGTCAACACAGACACAGAA 450  
QY 428 CGCGCGTTCTACGTCACATATGCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 487  
DB 451 CGCGCGTTCTACGTCACATATGCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 510  
QY 488 TAAGCGGGCATCAGTTTGGAGAACTACTCTTCAAGATTTTCTTACATCCCGGATGAAGAG 547  
DB 511 TAAGCGGGCATCAGTTTGGAGAACTACTCTTCAAGATTTTCTTACATCCCGGATGAAGAG 570  
QY 548 TGAGTCCCTTTCGCCCCCTCAGCGAGCCAGCGTGGGAGACCACTCTTCCCGGAGCAAG 607  
DB 571 TGAGTCCCTTTCGCCCCCTCAGCGAGCCAGCGTGGGAGACCACTCTTCCCGGAGCAAG 630  
QY 608 GCCACGCCCTGGGGGACCTTCTCAGGCCAGACAGATTTGATTTCCCGCTGCGGATCCTGG 667  
DB 631 GCCACGCCCTGGGGGACCTTCTCAGGCCAGACAGATTTGATTTCCCGCTGCGGATCCTGG 690  
QY 668 TCCCCACCCAGTTTGTGTGGTGCATCATCGGAAAGAGGAGGCTTGACCATAAAGAAATCA 727  
DB 691 TCCCCACCCAGTTTGTGTGGTGCATCATCGGAAAGAGGAGGCTTGACCATAAAGAAATCA 750  
QY 728 CTAAGCAGACCCAGTCCCGGTTAGATATCCATAGAAAAGAGAACTCTGGAGCTGCGAGAG 787  
DB 751 CTAAGCAGACCCAGTCCCGGTTAGATATCCATAGAAAAGAGAACTCTGGAGCTGCGAGAG 810  
QY 788 AGCCTGTCCATCCATCCATCCAGCCAGAGGGGACTTCTGAAGCATGCCGATCATTTCTTG 847  
DB 811 AGCCTGTCCATCCATCCAGCCAGAGGGGACTTCTGAAGCATGCCGATCATTTCTTG 870  
QY 848 AATATCATGAGAAAGAGGAGATGAGACCAAACTAGCCGAGAGATTCCTCTGAAAATCT 907  
DB 871 AATATCATGAGAAAGAGGAGATGAGACCAAACTAGCCGAGAGATTCCTCTGAAAATCT 930  
QY 908 TGCCACATAGTGGCTTGGTGGAGACTGATGGAAAAGAGGAGAGAAATTTGAAGAAA 967  
DB 931 TGCCACATAGTGGCTTGGTGGAGACTGATGGAAAAGAGGAGAGAAATTTGAAGAAA 990  
QY 968 TTCAATGAAACAGGAGCCAGATACATCTATCTTTTGCAGGATTTTCAGAGATATACA 1027  
DB 991 TTGAAATGAAACAGGAGCCAGATACATCTATCTTTTGCAGGATTTTCAGAGATATACA 1050  
QY 1028 ACCCGAAAGAAACCATCTGCTGAAAGGCGACAGTTGAGGCGCTGTGCCAGTGTGAGATAG 1087  
DB 1051 ACCCGAAAGAAACCATCTGCTGAAAGGCGACAGTTGAGGCGCTGTGCCAGTGTGAGATAG 1110  
QY 1088 AGATTATGAAAGAGCTGCTGAGGCGCTTTGAAAATGATATGCTGGCTGTAAACCAAG 1147  
DB 1111 AGATTATGAAAGAGCTGCTGAGGCGCTTTGAAAATGATATGCTGGCTGTAAACCAAG 1170  
QY 1148 CCAATCTGATCCAGAGGTTGAACCTCAGGSCACTTGGSCATCTTTTCAACAGAGACTGTCOG 1207  
DB 1171 CCAATCTGATCCAGAGGTTGAACCTCAGGSCACTTGGSCATCTTTTCAACAGAGACTGTCOG 1230  
QY 1208 TGCTATCTCCACAGAGGCGCCCGGAGAGCTCCCGCGCTGCCCTTACACCCCTTCA 1267  
DB 1231 TGCTATCTCCACAGAGGCGCCCGGAGAGCTCCCGCGCTGCCCTTACACCCCTTCA 1290  
QY 1268 CTAACCACTCCGAGTACTTCTCAGGCTGTACCCCATACACAGTTTGGCGCGTTCOGC 1327  
DB 1291 CTAACCACTCCGAGTACTTCTCAGGCTGTACCCCATACACAGTTTGGCGCGTTCOGC 1350  
QY 1328 ATCATCACTCTTATCCAGAGCAGAGATTTGTGAATCTTCTTCAACCAACCCAGGCTGTGG 1387  
DB 1351 ATCATCACTCTTATCCAGAGCAGAGATTTGTGAATCTTCTTCAACCAACCCAGGCTGTGG 1410  
QY 1388 GCGCCATCATCGGAAAGAGGGGCGACACATCAACAGCTGGCGAGATTCGCCGAGCCT 1447  
DB 1411 GCGCCATCATCGGAAAGAGGGGCGACACATCAACAGCTGGCGAGATTCGCCGAGCCT 1470



QY	1448	CTATCAAGATTGCCCTCGGGAAGCCAGAGCTCAGCGAAGAGTGTGTCATCATCACCG	1507
Db	1471	CTATCAAGATTGCCCTCGGGAAGCCAGAGCTCAGCGAAGAGTGTGTCATCATCACCG	1530
QY	1508	GGCCACCGGAGCCAGATTCAAGGCCCGAGGACGGATCTTTGGGAAAATTGAAGAGGAAA	1567
Db	1531	GGCCACCGGAGCCAGATTCAAGGCCCGAGGACGGATCTTTGGGAAAATTGAAGAGGAAA	1590
QY	1568	ACTTCTTTAAACCCCAAGGAAGTGAAGCTGGAGCGCATACAGTGCCTCTTCCA	1627
Db	1591	ACTTCTTTAAACCCCAAGGAAGTGAAGCTGGAGCGCATACAGTGCCTCTTCCA	1650
QY	1628	CAGCTGGCCGGTGATTTGGCAAGGTGGCAACCGCTGAAACGAACTGCAGAACTTAAACA	1687
Db	1651	CAGCTGGCCGGTGATTTGGCAAGGTGGCAACCGCTGAAACGAACTGCAGAACTTAAACA	1710
QY	1688	GTGCAGAAGTCATCTGTGCTCTGTGACCAAAACCCAGATGAAATGAGGAAGTATCGTCA	1747
Db	1711	GTGCAGAAGTCATCTGTGCTCTGTGACCAAAACCCAGATGAAATGAGGAAGTATCGTCA	1770
QY	1748	GAATTTATCGGCGACTTCTTTGTAGCCAGCTGCGACAGCGCAAGATCAGGGAATTTGTAC	1807
Db	1771	GAATTTATCGGCGACTTCTTTGTAGCCAGCTGCGACAGCGCAAGATCAGGGAATTTGTAC	1830
QY	1808	AACAGGTGAAGCAGCAGGACGAGAAATACCCCTCAGGAGTGCCTTCACAGCGCAGCAAGT	1867
Db	1831	AACAGGTGAAGCAGCAGGACGAGAAATACCCCTCAGGAGTGCCTTCACAGCGCAGCAAGT	1890
QY	1868	GAGGCTCCACAGGCACAGCAAAACCAAGGATGAATGTAGCCCTTCCAAACCTGACAG	1927
Db	1891	GAGGCTCCACAGGCACAGCAAAACCAAGGATGAATGTAGCCCTTCCAAACCTGACAG	1950
QY	1928	AATGAGACAAACGCGAGCCAGATCGGAGCAAAACCAAGACCATCTGAGGAATGAG	1987
Db	1951	AATGAGACAAACGCGAGCCAGATCGGAGCAAAACCAAGACCATCTGAGGAATGAG	2010
QY	1988	AAGTCTCGGAGCGCGCGAGGACTCTGCGAGGCCCTTGAGAACCCCGGCGCGAGGAG	2047
Db	2011	AAGTCTCGGAGCGCGCGAGGACTCTGCGAGGCCCTTGAGAACCCCGGCGCGAGGAG	2070
QY	2048	GGGCGGGGAAGTCAAGCAGGTTTGCAGAACCCAGCCAGCCCGCTCCCGCCCGCCAGG	2107
Db	2071	GGGCGGGGAAGTCAAGCAGGTTTGCAGAACCCAGCCAGCCCGCTCCCGCCCGCCAGG	2130
QY	2108	GCTTCTGAGGCTTCAGGCCATCCACTTCACCATCCACTCGGATCTCTGAACTCCAC	2167
Db	2131	GCTTCTGAGGCTTCAGGCCATCCACTTCACCATCCACTCGGATCTCTGAACTCCAC	2190
QY	2168	GACGCTATCCCTTTTGTAGTTGAACCTAACATAGGTGAACGTTTCAAAGCCAAAGATGC	2227
Db	2191	GACGCTATCCCTTTTGTAGTTGAACCTAACATAGGTGAACGTTTCAAAGCCAAAGATGC	2250
QY	2228	ACACCTTTTCTGTGGCAAAATCGTCTGTACATGTGTGTACATATTAGAAAGGGAGA	2287
Db	2251	ACACCTTTTCTGTGGCAAAATCGTCTGTACATGTGTGTACATATTAGAAAGGGAGA	2310
QY	2288	TGTTAAGATATGTGCCCTGTGGTTTACACAGGGTGCCTGCAGCGGTATATATTAGAA	2347
Db	2311	TGTTAAGATATGTGCCCTGTGGTTTACACAGGGTGCCTGCAGCGGTATATATTAGAA	2370
QY	2348	ATATATATCAATAACTCAACTCAATTTTAAATCAATTAATTTTCTTCT	2407
Db	2371	ATATATATCAATAACTCAACTCAATTTTAAATCAATTAATTTTCTTCT	2430
QY	2408	TTTTAAAGAGAAAGCAGGCTTTTCTAGACTTTTAAAGATAAAGTCTTTGGGAGGTCTCAC	2467
Db	2431	TTTTAAAGAGAAAGCAGGCTTTTCTAGACTTTTAAAGATAAAGTCTTTGGGAGGTCTCAC	2490
QY	2468	GGTGTAGAGAGGCTTTGAGGCCACCGCAAAATTTCAACCCAGAGGGAATCTCGTCG	2527
Db	2491	GGTGTAGAGAGGCTTTGAGGCCACCGCAAAATTTCAACCCAGAGGGAATCTCGTCG	2550
QY	2528	GAAGGACACTCAGGCAGTTCTGGATCACTGTGTATGTCAACAGAGGGATACCGTCTC	2587
Db	2551	GAAGGACACTCAGGCAGTTCTGGATCACTGTGTATGTCAACAGAGGGATACCGTCTC	2610
QY	2588	CTTGAAGAGGAAACTCTGTCTACTCTCATGCTGTGTAGCTCATACACCAATTTCTCTTT	2647
Db	2611	CTTGAAGAGGAAACTCTGTCTACTCTCATGCTGTGTAGCTCATACACCAATTTCTCTTT	2670
QY	2648	GCTTCACAGGTTTAAACCTGTTTTTTCATACTGCTATATAATTTCTGTCTCTCTCTG	2707
Db	2671	GCTTCACAGGTTTAAACCTGTTTTTTCATACTGCTATATAATTTCTGTCTCTCTCTG	2730
QY	2708	TTTATCTCTCCCTCCCTCCCTCCCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	2767
Db	2731	TTTATCTCTCCCTCCCTCCCTCCCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	2790
QY	2768	ATCCCTCCATCTCAATCCGCTATCTAAGCA--CCGCCCCCCCCCCCCCAGCAAGAGTGT	2825
Db	2791	ATCCCTCCATCTCAATCCGCTATCTAAGCA--CCGCCCCCCCCCCCCCAGCAAGAGTGT	2850
QY	2826	CTGAGTATCACATCACAAAGAAACAAAGCGAAACACACAAACCAAGCCTCAACTTAC	2885
Db	2851	CTGAGTATCACATCACAAAGAAACAAAGCGAAACACACAAACCAAGCCTCAACTTAC	2910
QY	2886	ACTTGGTTACTCAAAAGAACCAAGAGTCAATGGTACTTGTCTCTAGCGTTTTTGGAGAGAA	2945
Db	2911	ACTTGGTTACTCAAAAGAACCAAGAGTCAATGGTACTTGTCTCTAGCGTTTTTGGAGAGAA	2970
QY	2946	AACAGGAACCCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA	3005
Db	2971	AACAGGAACCCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA	3030
QY	3006	ATGTAATTTTGTCTTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTT	3065
Db	3031	ATGTAATTTTGTCTTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTT	3090
QY	3066	TTTCTTT--AAAAAATAAATGTGAGGAAAGTAGAAATTTTACCAGGTTTGTGGCCGAGG	3124
Db	3091	TTTCTTTAAAAAATAAATGTGAGGAAAGTAGAAATTTTACCAGGTTTGTGGCCGAGG	3150
QY	3125	GCGTTAAATTCACAGATTTTTTAAACGAGAAACACACAGAAAGAGTACTCAGGTGT	3184
Db	3151	GCGTTAAATTCACAGATTTTTTAAACGAGAAACACACAGAAAGAGTACTCAGGTGT	3210
QY	3185	TTTTTACCTCAGACCTTGTCTTGTCTTGTCTTGTCTTGTCTTGTCTTGTCTTGTCTTGT	3244
Db	3211	TTTTTACCTCAGACCTTGTCTTGTCTTGTCTTGTCTTGTCTTGTCTTGTCTTGTCTTGT	3270
QY	3245	AGCATTTTTTTTATTTTAAATAAATAGTTTGGAAAAAATAAAGATATCAACTGCCA	3304
Db	3271	AGCATTTTTTTTATTTTAAATAAATAGTTTGGAAAAAATAAAGATATCAACTGCCA	3330
QY	3305	GCTTGGAGAGGTGACAGTCCAGTGTGCAACAGCTGTTCTGAATTTCTTCCGCTAGCC	3364
Db	3331	GCTTGGAGAGGTGACAGTCCAGTGTGCAACAGCTGTTCTGAATTTCTTCCGCTAGCC	3390
QY	3365	AAGAACCC 3371	
Db	3391	AAGAACCC 3397	
RESULT 4			
ID	AAZ36154	standard; DNA; 3283 BP.	
XX	AAZ36154		
AC	AAZ36154;		
XX			
DT	11-FEB-2000	(first entry)	
XX			
DE	An alternative form of DNA encoding cancer associated antigen KOC-3.		
XX	Cancer associated antigen; KOC-3; cancer; vaccine; CTI7; ss.		
XX	Homo sapiens.		
OS			



[illegible]

Db 2701 GTATCACATCACAAAAGGAAACAAAGCGAAACACACAAACAGCCTCACTTACACTT 2760  
QY 2890 GGTACTCAAAAGACACAGAGTCAATGGTACTTGTCTAGCGTTTGGAGAGGAAACA 2949  
Db 2761 GGTACTCAAAAGACACAGAGTCAATGGTACTTGTCTAGCGTTTGGAGAGGAAACA 2820  
QY 2950 GGAACCCACCAACCAACCAATCAACCAACCAAGAAAGAAATTCACATGAAGAAATGT 3009  
Db 2821 GGAACCCACCAACCAACCAATCAACCAACCAAGAAAGAAATTCACATGAAGAAATGT 2880  
QY 3010 ATTTGTCTTTTGGTATTTGGTATTAAGCCATCAATATTCAGCAAAATGATTCCTTTC 3069  
Db 2881 ATTTGTCTTTTGGTATTTGGTATTAAGCCATCAATATTCAGCAAAATGATTCCTTTC 2940  
QY 3070 TTTAAAAAATAAATGCGAGCAAGATGAATTTACCAAGTTGTGGCCAGGCGTT 3129  
Db 2941 TTTAAAAAATAAATGCGAGCAAGATGAATTTACCAAGTTGTGGCCAGGCGTT 3000  
QY 3130 AAATTCACAGATTTTTTTAAACGAGAAACACACAGAAAGCTACCTCAGGTGTTTTTA 3189  
Db 3001 AAATTCACAGATTTTTTTAAACGAGAAACACACAGAAAGCTACCTCAGGTGTTTTTA 3060  
QY 3190 CCTAGACCTTGTCTTGTGTTTCCCTTAGAGATTTTGAAGCTGTAGTTGAGCAT 3249  
Db 3061 CCTAGACCTTGTCTTGTGTTTCCCTTAGAGATTTTGAAGCTGTAGTTGAGCAT 3120  
QY 3250 TTTTATTTTATTTTAAATGAATGTTGGAAAGAAACACACAGAAAGCTACCTCAGGTGTTTTTA 3309  
Db 3121 TTTTATTTTATTTTAAATGAATGTTGGAAAGAAACACACAGAAAGCTACCTCAGGTGTTTTTA 3180  
QY 3310 GAGAAAGTGACAGTCCAAAGTGTGCAACAGCTGTTCTGAAATGTTCTTCGCTAGCCAGAA 3369  
Db 3181 GAGAAAGTGACAGTCCAAAGTGTGCAACAGCTGTTCTGAAATGTTCTTCGCTAGCCAGAA 3240  
QY 3370 CNATATGCGCTCTTTTGGACAAACCTTGAAATGTTTATTT 3412  
Db 3241 CNATATGCGCTCTTTTGGACAAACCTTGAAATGTTTATTT 3283

RESULT 5

AAS70981  
ID AAS70981 standard; cDNA; 2010 BP.  
AC AAS70981;  
XX  
XX  
DT 13-FEB-2002 (first entry)  
XX  
DE DNA encoding novel human diagnostic protein #6785.  
XX  
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;  
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.  
XX  
XX Homo sapiens.  
XX  
XX WO200175067-A2.  
XX  
XX 11-OCT-2001.  
XX  
XX 30-MAR-2001; 2001WO-US008631.  
XX  
XX 31-MAR-2000; 2000US-00540217.  
PR 23-AUG-2000; 2000US-00649167.  
XX  
XX (HYSE-) HYSEQ INC.  
XX  
XX Drmanac RT, Liu C, Tang YT;  
XX  
XX WPI; 2001-639362/73.  
DR P-ESDB; A5G06794.  
XX  
XX New isolated polynucleotide and encoded polypeptides, useful in  
PT diagnostics, forensics, gene mapping, identification of mutations

PT responsible for genetic disorders or other traits and to assess  
PT biodiversity.  
XX  
PS Claim 1; SEQ ID NO 6785; 103pp; English.  
XX  
CC The invention relates to isolated polynucleotide (I) and polypeptide (II)  
CC sequences. (I) is useful as hybridisation probes, polymerase chain  
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
CC and in recombinant production of (II). The polynucleotides are also used  
CC in diagnostics as expressed sequence tags for identifying expressed  
CC genes. (I) is useful in gene therapy techniques to restore normal  
CC activity of (II) or to treat disease states involving (II). (II) is  
CC useful for generating antibodies against it, detecting or quantitating a  
CC polypeptide in tissue, as molecular weight markers and as a food  
CC supplement. (III) and its binding partners are useful in medical imaging  
CC of sites expressing (II). (I) and (II) are useful for treating disorders  
CC involving aberrant protein expression or biological activity. The  
CC polypeptide and polynucleotide sequences have applications in  
CC diagnostics, forensics, gene mapping, identification of mutations  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic  
CC coding sequences of the invention. Note: The sequence data for this  
CC patent did not appear in the printed specification, but was obtained in  
CC electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX

SQ Sequence 2010 BP; 565 A; 552 C; 532 G; 361 T; 0 U; 0 Other;

Query Match 57.9%; Score 1977; DB 5; Length 2010;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1977; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 57 GGGGAAGAGACGATGATGAACAGCTTTTACATCGGAAACCTGAGCCCGCGTCACCGC 116  
Db 34 GGGGAAGAGACGATGATGAACAGCTTTTACATCGGAAACCTGAGCCCGCGTCACCGC 93  
QY 117 CGAGACCTCCGCGAGCTCTTTGGGACAGGAAGCTGCCCTGGGACAGAGTCTCTCT 176  
Db 94 CGAGACCTCCGCGAGCTCTTTGGGACAGGAAGCTGCCCTGGGACAGAGTCTCTCT 153  
QY 177 GAAGTCCGGCTACGCTTCTGTGGACTACCCGACACAGAACTGGGCACTCCGCGCATCGA 236  
Db 154 GAAGTCCGGCTACGCTTCTGTGGACTACCCGACACAGAACTGGGCACTCCGCGCATCGA 213  
QY 237 GACCTCTCGGTTAAAGTGAATTCATCGGAAATCATGGAAGTTGATTACTCAGTCTC 296  
Db 214 GACCTCTCGGTTAAAGTGAATTCATCGGAAATCATGGAAGTTGATTACTCAGTCTC 273  
QY 297 TAAAAAGCTAAGGAGCAGGAAATTCAGATTCGAAACATCCCTCCTCAGCTGCGAGTGGGA 356  
Db 274 TAAAAAGCTAAGGAGCAGGAAATTCAGATTCGAAACATCCCTCCTCAGCTGCGAGTGGGA 333  
QY 357 GGTGTGGATGGACTTTTGGCTCAATATGGGACAGTGGAGAATGTGGAACAAAGTCAACAC 416  
Db 334 GGTGTGGATGGACTTTTGGCTCAATATGGGACAGTGGAGAATGTGGAACAAAGTCAACAC 393  
QY 417 AGACACAGAAACCGCGTTGTCAACGTCATATGCAACAGAGAAAGCAAAATAGC 476  
Db 394 AGACACAGAAACCGCGTTGTCAACGTCATATGCAACAGAGAAAGCAAAATAGC 453  
QY 477 CATGGAGAAGCTAAGCGGGCATCAGTTTGGAACTACTCTTCAAGATTTCTTACATCCC 536  
Db 454 CATGGAGAAGCTAAGCGGGCATCAGTTTGGAACTACTCTTCAAGATTTCTTACATCCC 513  
QY 537 GGATGAAGAGTGAGTCCCTTCGCCCCCTCAGCGAGCCAGCGTGGGGACCACTCTTC 596  
Db 514 GGATGAAGAGTGAGTCCCTTCGCCCCCTCAGCGAGCCAGCGTGGGGACCACTCTTC 573  
QY 597 CCGGAGCAAGGCCACCGCCCTCGGGGCATCTTCTCAGGCCAGACAGATTCGATTTCCCGCT 656  
Db 574 CCGGAGCAAGGCCACCGCCCTCGGGGCATCTTCTCAGGCCAGACAGATTCGATTTCCCGCT 633

QY 657 GCGGATCTTGTCCTCCACCCAGTTTGTGTGTCATCATCGGAAAGAGGGCTTGACCAT 716  
Db 634 GCGGATCTTGTCCTCCACCCAGTTTGTGTGTCATCATCGGAAAGAGGGCTTGACCAT 693  
QY 717 AAAGAAATCATCACTAAGCAGACCCAGTCCCGGGTAGATATCCATAGAAAAAGAACTCTGG 776  
Db 694 AAAGAAATCATCACTAAGCAGACCCAGTCCCGGGTAGATATCCATAGAAAAAGAACTCTGG 753  
QY 777 AGCTGAGAGAACTCTGTCAATCCATGCCACCCAGAGGGGACTTCTGAAGCATGCCG 836  
Db 754 AGCTGAGAGAACTCTGTCAATCCATGCCACCCAGAGGGGACTTCTGAAGCATGCCG 813  
QY 837 CATGATCTTGAATCATGTCAGAAAGAGGCAGATGAGACCAACTAGCCGAGAGATTCC 896  
Db 814 CATGATCTTGAATCATGTCAGAAAGAGGCAGATGAGACCAACTAGCCGAGAGATTCC 873  
QY 897 TCTGAAAAATCTTGGCACAACAATGGCTTGGTGGAGACTGATTGGAAAAAGAGCGAGAA 956  
Db 874 TCTGAAAAATCTTGGCACAACAATGGCTTGGTGGAGACTGATTGGAAAAAGAGCGAGAA 933  
QY 957 TTTGAAAAAATTTGAAATGAACAGGACCAAGATAACAATCTCATCTTTTGAGGATTT 1016  
Db 934 TTTGAAAAAATTTGAAATGAACAGGACCAAGATAACAATCTCATCTTTTGAGGATTT 993  
QY 1017 GAGCATATACAAACCGGAAAGAACCATCACTGTGAAGGGCACAGTTGAGGCCCTGTGCCAG 1076  
Db 994 GAGCATATACAAACCGGAAAGAACCATCACTGTGAAGGGCACAGTTGAGGCCCTGTGCCAG 1053  
QY 1077 TGTGTGATAGATGATATGAAGAGCTGCGTGAAGCCCTTTGAAAAATGATATGCTGGCTGT 1136  
Db 1054 TGTGTGATAGATGATATGAAGAGCTGCGTGAAGCCCTTTGAAAAATGATATGCTGGCTGT 1113  
QY 1137 TAACCAACAAAGCCAAATCTGATCCAGGGTTGAACCTCAGCGCACTTGGCATCTTTCAAC 1196  
Db 1114 TAACCAACAAAGCCAAATCTGATCCAGGGTTGAACCTCAGCGCACTTGGCATCTTTCAAC 1173  
QY 1197 AGGACTGTGCGTGTATCTCCACAGCAGGCGCCGCGAGCTCCCGCGCTGCCCTTA 1256  
Db 1174 AGGACTGTGCGTGTATCTCCACAGCAGGCGCCGCGAGCTCCCGCGCTGCCCTTA 1233  
QY 1257 CCACCCCTTCACTACCCACTCCGATATCTTCAGAGCTGTACCCCATCACCAGTTTGG 1316  
Db 1234 CCACCCCTTCACTACCCACTCCGATATCTTCAGAGCTGTACCCCATCACCAGTTTGG 1293  
QY 1317 CCGTTCCTCCGATCATCACTCTTATCCAGAGCAGGAGATTGTGAATCTCTTCATCCCAAC 1376  
Db 1294 CCGTTCCTCCGATCATCACTCTTATCCAGAGCAGGAGATTGTGAATCTCTTCATCCCAAC 1353  
QY 1377 CCAGGCTGTGGGCCATCATCGGAGAGAGGGGGCACATCAACAGCTGGCGAGATT 1436  
Db 1354 CCAGGCTGTGGGCCATCATCGGAGAGAGGGGGCACATCAACAGCTGGCGAGATT 1413  
QY 1437 CGCGGAGCCTCTATCAAGATTGCCCTTGGGAAAGGCCAGACTCAGCGAAAGAGTGGT 1496  
Db 1414 CGCGGAGCCTCTATCAAGATTGCCCTTGGGAAAGGCCAGACTCAGCGAAAGAGTGGT 1473  
QY 1497 CATCATCCGGGCCACCGGAAGCCGATTCAGGCCCCAGGACGATCTTTGGGAAACT 1556  
Db 1474 CATCATCCGGGCCACCGGAAGCCGATTCAGGCCCCAGGACGATCTTTGGGAAACT 1533  
QY 1557 GAAAGAGAAACTTTCTTAACCCCAAGAGAGAGTGAAGCTGGAAGCGCATATCAGAGT 1616  
Db 1534 GAAAGAGAAACTTTCTTAACCCCAAGAGAGAGTGAAGCTGGAAGCGCATATCAGAGT 1593  
QY 1617 GCCCTCTTCCACAGCTGGCCGGTGAATGGCAAGGTGGCAAGCCGTGAACGAACCTGCA 1676  
Db 1594 GCCCTCTTCCACAGCTGGCCGGTGAATGGCAAGGTGGCAAGCCGTGAACGAACCTGCA 1653  
QY 1677 GAACTTAACAGTGCAGAGTCACTGTCCTGTCGACCAAAAGCCAGATGAAATAGGA 1736  
Db 1654 GAACTTAACAGTGCAGAGTCACTGTCCTGTCGACCAAAAGCCAGATGAAATAGGA 1713  
QY 1737 AGTGATCGTCAGAAATATCGGGCCTCTTTGCTAGCCAGACTGCACAGCCAGATCAG 1796

Db 1714 AGTGATCGTCAGAAATATCGGCACCTTCTTTGCTAGCCAGACTGCACAGCGCAAGATCAG 1773  
QY 1797 GGAATTTGTACAACAGGTGAAGCAGCAGGAGCAGAAATACCTCTCAGGAGTGCCTCACA 1856  
Db 1774 GGAATTTGTACAACAGGTGAAGCAGCAGGAGCAGAAATACCTCTCAGGAGTGCCTCACA 1833  
QY 1857 GGCACAAAGTGAAGTCCCAAGGACCAAGCAAAACAACGGATGAATGTAGCCCTTCCA 1916  
Db 1834 GGCACAAAGTGAAGTCCCAAGGACCAAGCAAAACAACGGATGAATGTAGCCCTTCCA 1893  
QY 1917 ACACCTTGACAGATGAGACCAACAGCAGCCAGCAGATCGGAGCAAAACCAAGACCATC 1976  
Db 1894 ACACCTTGACAGATGAGACCAACAGCAGCCAGCAGATCGGAGCAAAACCAAGACCATC 1953  
QY 1977 TGAGGAATGAGAGTCTGCGGAGCGCCAGGAGCTCTGCGAGGCCCTGAGAACCC 2033  
Db 1954 TGAGGAATGAGAGTCTGCGGAGCGCCAGGAGCTCTGCGAGGCCCTGAGAACCC 2010

RESULT 6  
AAS70982  
ID AAS70982 standard; cDNA; 3667 BP.  
XX  
AC AAS70982;  
XX DT 13-FEB-2002 (first entry)  
XX DE DNA encoding novel human diagnostic protein #6786.  
XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
XX KW food supplement; medical imaging; diagnostic; genetic disorder; ss.  
XX OS Homo sapiens.  
XX PN WO200175067-A2.  
XX PD 11-OCT-2001.  
XX PF 30-MAR-2001; 2001WO-US008631.  
XX PR 31-MAR-2000; 2000US-00540217.  
XX PR 23-AUG-2000; 2000US-00649167.  
XX PA (HYSE-) HYSEQ INC.  
XX PI Drmanac RT, Liu C, Tang YT;  
XX PI WPI; 2001-639362/73.  
XX DR P-PSDB; ABG06795.  
XX PT New isolated polynucleotide and encoded polypeptides, useful in  
PT diagnostics, forensics, gene mapping, identification of mutations  
PT responsible for genetic disorders or other traits and to assess  
PT biodiversity.  
XX  
PS Claim 1; SEQ ID NO 6786; 103pp; English.  
XX  
CC The invention relates to isolated polynucleotide (I) and polypeptide (II)  
CC sequences. (I) is useful as hybridisation probes, polymerase chain  
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
CC and in recombinant production of (II). The polynucleotides are also used  
CC in diagnostics as expressed sequence tags for identifying expressed  
CC genes. (I) is useful in gene therapy techniques to restore normal  
CC activity of (II) or to treat disease states involving (II). (II) is  
CC useful for generating antibodies against it, detecting or quantitating a  
CC polypeptide in tissue, as molecular weight markers and as a food  
CC supplement. (II) and its binding partners are useful in medical imaging  
CC of sites expressing (II). (I) and (II) are useful for treating disorders  
CC involving aberrant protein expression or biological activity. The  
CC polypeptide and polynucleotide sequences have applications in  
CC diagnostics, forensics, gene mapping, identification of mutations  
CC responsible for genetic disorders or other traits to assess biodiversity

CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic  
CC coding sequences of the invention. Note: The sequence data for this  
CC patent did not appear in the printed specification, but was obtained in  
CC electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 3667 BP; 1081 A; 882 C; 846 G; 858 T; 0 U; 0 Other;

Query Match 53.0%; Score 1810; DB 5; Length 3667;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 2100; Conservative 0; Mismatches 1; Indels 2; Gaps 2;  
QY 1270 ACCCACTCGGATCTTCTCCAGCTGTACCCCATCACAGTTTGGCCGTTCCGGCAT 1329  
DB 1507 ACCCACTCGGATCTTCTCCAGCTGTACCCCATCACAGTTTGGCCGTTCCGGCAT 1566  
QY 1330 CATCATCTTATCCAGAGCAGGAGTTGTGAATCTTTCATCCCAACCCAGCTGTGGC 1389  
DB 1567 CATCATCTTATCCAGAGCAGGAGTTGTGAATCTTTCATCCCAACCCAGCTGTGGC 1626  
QY 1390 GCCATCCTGGGAAGAGGGGCACACATCAACAGCTGGCGAGATTGGCCGAGCCTCT 1449  
DB 1627 GCCATCCTGGGAAGAGGGGCACACATCAACAGCTGGCGAGATTGGCCGAGCCTCT 1686  
QY 1450 ATCAAGATTGCCCTTGGGAAAGGGCCAGACGTCAGCGAAAGGATGGTCAATCACCGGG 1509  
DB 1687 ATCAAGATTGCCCTTGGGAAAGGGCCAGACGTCAGCGAAAGGATGGTCAATCACCGGG 1746  
QY 1510 CCACCGAAGCCAGTCTCAAGGCCCAGGAGCGGATCTTTGGAAACTGAAAGAGAAAC 1569  
DB 1747 CCACCGAAGCCAGTCTCAAGGCCCAGGAGCGGATCTTTGGAAACTGAAAGAGAAAC 1806  
QY 1570 TTTCTTTAAACCAAGAGAGTGAAGCTGGAGCGCATACAGATGCCCTCTTCCACA 1629  
DB 1807 TTTCTTTAAACCAAGAGTGAAGCTGGAGCGCATACAGATGCCCTCTTCCACA 1866  
QY 1630 GCTGCGCGGATTTGGCAAGGTTGGCAAGCCGTGAACGAACTCAGAACTTAAACAGT 1689  
DB 1867 GCTGCGCGGATTTGGCAAGGTTGGCAAGCCGTGAACGAACTCAGAACTTAAACAGT 1926  
QY 1690 GCAGAGTCATCGTCCCTCGTGACCAAGCGCAGATGAAGATGAGAAATGTGACAA 1749  
DB 1927 GCAGAGTCATCGTCCCTCGTGACCAAGCGCAGATGAAGATGAGAAATGTGACAA 1986  
QY 1750 ATTATCGGCACTCTTTGTAGCCAGACTGCACAGCGCAAGATCAGGAAATGTGACAA 1809  
DB 1987 ATTATCGGCACTCTTTGTAGCCAGACTGCACAGCGCAAGATCAGGAAATGTGACAA 2046  
QY 1810 CAGGTGAAGCAGCAGGAGCAAAATPACCTCAGGAGTCGCTCAGCGCAGCAAGTGA 1869  
DB 2047 CAGGTGAAGCAGCAGGAGCAAAATPACCTCAGGAGTCGCTCAGCGCAGCAAGTGA 2106  
QY 1870 GGCTCCCAAGCAGCAGCAAAACAAAGGATGAATGTAGCCCTTCCAACTGACAGAA 1929  
DB 2107 GGCTCCCAAGCAGCAGCAAAACAAAGGATGAATGTAGCCCTTCCAACTGACAGAA 2166  
QY 1930 TGAGACCAAAACGACCCAGCAGATCGGAGCAAAACCAAGACCATCTGAGGAATGAA 1989  
DB 2167 TGAGACCAAAACGACCCAGCAGATCGGAGCAAAACCAAGACCATCTGAGGAATGAA 2226  
QY 1990 GTCTGCGAGCGCGCAGGAGCTCTGCGAGGCCCTGAGAACCCAGGGCCGAGGAGG 2049  
DB 2227 GTCTGCGAGCGCGCAGGAGCTCTGCGAGGCCCTGAGAACCCAGGGCCGAGGAGG 2286  
QY 2050 GCGGGAGGTCAGCCAGTTTGCAGAACACAGCCAGCCCGCTCCGCCCCCGAGGGC 2109  
DB 2287 GCGGGAGGTCAGCCAGTTTGCAGAACACAGCCAGCCCGCTCCGCCCCCGAGGGC 2346  
QY 2110 TTCTGCGAGGTCAGCCATCCACTTCAACATCCACTCGGATCTCTCTGAACTCCACGA 2169  
DB 2347 TTCTGCGAGGTCAGCCATCCACTTCAACATCCACTCGGATCTCTCTGAACTCCACGA 2406

QY 2170 CGCTATCCCTTTTACCTTGAACCTAACTAGCTGAACTGTTCAAAGCCAGCAAAATGCAC 2229  
DB 2407 CGCTATCCCTTTTACCTTGAACCTAACTAGCTGAACTGTTCAAAGCCAGCAAAATGCAC 2466  
QY 2230 ACCCTTTTCTGTGGCAAAATCGTCTGTACATGTGTGTACATATTTAGAAAGGAGATG 2289  
DB 2467 ACCCTTTTCTGTGGCAAAATCGTCTGTGTACATGTGTGTACATATTTAGAAAGGAGATG 2526  
QY 2290 TTAAGATATGCGCTGTGGGTTTACACAGGGTGCCTGACGGGTATATATTTTGAAT 2349  
DB 2527 TTAAGATATGCGCTGTGGGTTTACACAGGGTGCCTGACGGGTATATATTTTGAAT 2586  
QY 2350 AATATATCAAAATAACTCAACTAACTCCAAATTTTAAATCAATTTAAATTTTCTTT 2409  
DB 2587 AATATATCAAAATAACTCAACTAACTCCAAATTTTAAATCAATTTAAATTTTCTTT 2646  
QY 2410 TTAAGAGAAACGAGGCTTTTCTAGACTTTAAAGAAATAAAGTCTTTGGAGGTCTCAAG 2469  
DB 2647 TTAAGAGAAACGAGGCTTTTCTAGACTTTAAAGAAATAAAGTCTTTGGAGGTCTCAAG 2706  
QY 2470 TGTAGAGAGGAGCTTTGAGGCCACCCGACAAAATTCACCCAGAGGAAATCTCTCGGA 2529  
DB 2707 TGTAGAGAGGAGCTTTGAGGCCACCCGACAAAATTCACCCAGAGGAAATCTCTCGGA 2766  
QY 2530 AGGACACTCACGGCAGTTTCTGATCACCTGTGTATGTCAACAGAGGATACCGTCTCT 2589  
DB 2767 AGGACACTCACGGCAGTTTCTGATCACCTGTGTATGTCAACAGAGGATACCGTCTCT 2826  
QY 2590 TGAAGAGGAAACTCTGTCACTCTCATGCTGTCTAGCTCATACACCCATTTCTCTTTC 2649  
DB 2827 TGAAGAGGAAACTCTGTCACTCTCATGCTGTCTAGCTCATACACCCATTTCTCTTTC 2886  
QY 2650 TTCACAGTTTAACTGGTTTTTGTGATCTGTCTATATATTTCTCTCTCTCTCTCT 2709  
DB 2887 TTCACAGTTTAACTGGTTTTTGTGATCTGTCTATATATTTCTCTCTCTCTCTCT 2946  
QY 2710 TATCTCTCCCTCCCTCCCTCCCTCTCTCTCATCTCCATCTCTTTTGAATTTCTCT 2769  
DB 2947 TATCTCTCCCTCCCTCCCTCCCTCTCTCTCATCTCTCTCTCTCTCTCTCTCTCT 3006  
QY 2770 CCCTCCATCTCAATCCCGTATCTACGCA-CCCCCCCCCCCCCAGGCAAGAGTGTCTGA 2829  
DB 3007 CCCTCCATCTCAATCCCGTATCTACGCA-CCCCCCCCCCCCCAGGCAAGAGTGTCTGA 3065  
QY 2830 GTATCATATCACAAAAGGAAACAAAGCGAAACACACAAACCCAGCTCACTTACACTT 2889  
DB 3066 GTATCATATCACAAAAGGAAACAAAGCGAAACACACAAACCCAGCTCACTTACACTT 3125  
QY 2890 GGTACTCAAAAGAAACAGAGTCAATGGTACTTCTGTAGCGTTTGGAGAGGAAACA 2949  
DB 3126 GGTACTCAAAAGAAACAGAGTCAATGGTACTTCTGTAGCGTTTGGAGAGGAAACA 3185  
QY 2950 GGAAACCAACCAACCAATCAACCAAAACAAAGAAAAATTCACAAATCAAGAAATGT 3009  
DB 3186 GGAAACCAACCAACCAATCAACCAAAACAAAGAAAAATTCACAAATCAAGAAATGT 3245  
QY 3010 ATTTTGTCTTTTGTGATTAAGCCATCAATATTCAGCAAAATCAATTTCTCTTTC 3069  
DB 3246 ATTTTGTCTTTTGTGATTAAGCCATCAATATTCAGCAAAATCAATTTCTCTTTC 3305  
QY 3070 TTT-AAAAAATAAGTGGAGGAAAGTAGAAATTTTACCAAGTTTGGCCCGAGGCGT 3128  
DB 3306 TTTAAAAAATAAGTGGAGGAAAGTAGAAATTTTACCAAGTTTGGCCCGAGGCGT 3365  
QY 3129 TAAATTCAGATTTTTTAAACGAGAAAAACACACAGAAAGCTACTCTCAGGTGTTTTT 3188  
DB 3366 TAAATTCAGATTTTTTAAACGAGAAAAACACACAGAAAGCTACTCTCAGGTGTTTTT 3425  
QY 3189 ACCTCAGCCTCTCTCTGTGTCTTCCCTAGAGATTTGTAAAGCTGTAGTTGGAGCA 3248  
DB 3426 ACCTCAGCCTCTCTGTGTCTTCCCTAGAGATTTGTAAAGCTGTAGTTGGAGCA 3485  
QY 3249 TTTTATTTTAAATAAATAAGTGGAGGAAATAAGATATCAACTGCCAGCT 3308

3486	TTTTTTTATTTTTTAAATAAAGTGTGGAATAATCAACTGCCAGCT	3545
3309	GGAGAAGTGACAGTCCAAGTGTGCAACAGCTGTCTGAAATTGTCTTCCGCTAGCCCAAGA	3368
3546	GGAGAAGTGACAGTCCAAGTGTGCAACAGCTGTCTGAAATTGTCTTCCGCTAGCCCAAGA	3605
3369	ACC 3371	
3606	ACC 3608	
RESULT 7		
ABS76442		
ID	ABS76442 standard; cDNA; 3667 BP.	
XX	ABS76442;	
XX	11-DEC-2002 (first entry)	
XX	cDNA encoding human ovarian cancer marker M452.	
XX	Human; ovarian cancer; marker; cancer; familial history; brain disorder;	
KW	central nervous system disorder; bacterial meningitis; viral meningitis;	
KW	Alzheimer's disease; Parkinson's disease; cerebral oedema; hydrocephalus;	
KW	brain herniation; inflammation; encephalitis; testicular disorder;	
KW	nontuberculous granulomatous orchitis; connective tissue disorder;	
KW	heart disorder; ischaemic heart disease; atherosclerosis; neoplasm;	
KW	histological type; carcinogenic; ovarian cancer marker; gene; ss.	
XX	Homo sapiens.	
OS	WO200271928-A2.	
PN	19-SEP-2002.	
PD		
XX	14-MAR-2002; 2002WO-US007826.	
PF		
XX	14-MAR-2001; 2001US-0276025P.	
PR	14-MAR-2001; 2001US-0276025P.	
PR	10-AUG-2001; 2001US-0311732P.	
PR	19-SEP-2001; 2001US-0323580P.	
PR	26-SEP-2001; 2001US-0324967P.	
PR	26-SEP-2001; 2001US-0325102P.	
PR	26-SEP-2001; 2001US-0325149P.	
XX	(MILL-) MILLENNIUM PHARM INC.	
PA	Monahan JE, Gannavarapu M, Hoersch S, Kamatkar S, Kovatis SG;	
PI	Meyers RE, Morrissey WP, Olandt PJ, Sen A, Vieby PO, Mills GB;	
PI	Bast RC, Lu K, Schmandt RE, Zhao X, Glatt K;	
XX	WPI; 2002-723277/78.	
DR	P-PSDB; ABG96346.	
DR		
XX		
PT	Assessing whether a patient is afflicted with ovarian cancer, useful in	
PT	assessing the stage or progression of the disease, comprises comparing	
PT	the expression level of a cancer marker in a sample from a patient and	
PT	from a non cancer patient.	
XX	Disclosure; Page 262-263; 481pp; English.	
PS		
XX	The present invention relates to a new method for assessing whether a	
CC	patient is afflicted with ovarian cancer. The method involves comparing	
CC	the expression level of a marker in a patient sample and the normal level	
CC	of expression of the marker in a control non-ovarian cancer sample, where	
CC	the marker is selected from 363 cancer markers described in the	
CC	specification. The method of the invention is useful in diagnosing or	
CC	characterising cancer, in detecting the presence of cancer as early as	
CC	possible, and the recurrence of ovarian cancer. The method may also be	
CC	particular use with patients having an enhanced risk of developing	
CC	ovarian cancer (e.g. patients having a familial history of ovarian	
CC	cancer). The cancer markers may be used in the management and treatment	

QY 2050 GCGGGAAGGTGAGCCAGGTTTGCCAGAACACCGAGCCCGCTCCGCCCCCAGG3C 2109  
Db 2287 GCGGGAAGGTGAGCCAGGTTTGCCAGAACACCGAGCCCGCTCCGCCCCCAGG3C 2346  
QY 2110 TTCTGAGGCTTCAGCCATCCACTTCAACATCCACTCGGATCTCTCTGAACTCCACGA 2169  
Db 2347 TTCTGAGGCTTCAGCCATCCACTTCAACATCCACTCGGATCTCTCTGAACTCCACGA 2406  
QY 2170 GCGTATCCCTTTAGTCTGAACCTAAACATAGTGTGTAAGCGTGTCAAGCGAAGAAATGCAC 2229  
Db 2407 GCGTATCCCTTTAGTCTGAACCTAAACATAGTGTGTAAGCGTGTCAAGCGAAGAAATGCAC 2466  
QY 2230 ACCCTTTTCTGTGGCAATCGCTCTGTACATGTGTGTACATATTAGAAAGGGAAGATG 2289  
Db 2467 ACCCTTTTCTGTGGCAATCGCTCTGTACATGTGTGTACATATTAGAAAGGGAAGATG 2526  
QY 2290 TTAAGATATGTGGCTGTGGGTTACACAGGTCCTGCAGCGGTAAATATTATTTAGAAAT 2349  
Db 2527 TTAAGATATGTGGCTGTGGGTTACACAGGTCCTGCAGCGGTAAATATTATTTAGAAAT 2586  
QY 2350 AATATATCAATTAACCTCAACTAACTCCAAATTTTAAATCAATTAATTTTCTTTTCTTT 2409  
Db 2587 AATATATCAATTAACCTCAACTAACTCCAAATTTTAAATCAATTAATTTTCTTTTCTTT 2646  
QY 2410 TTAAGAGAAAGCAGGCTTTTCTAGACTTTTAAAGATAAAGTCTTTGGGAGGTCTCACGG 2469  
Db 2647 TTAAGAGAAAGCAGGCTTTTCTAGACTTTTAAAGATAAAGTCTTTGGGAGGTCTCACGG 2706  
QY 2470 TGTAGAGAGGACTTTGAGGCCACCCGCACAAAATTCACCGAGGAAATCTGTCGGA 2529  
Db 2707 TGTAGAGAGGACTTTGAGGCCACCCGCACAAAATTCACCGAGGAAATCTGTCGGA 2766  
QY 2530 AGGACACTCAGGCAAGTCTTGATCACTGTGTATGTCAACAGAGGATACCGTCTCT 2589  
Db 2767 AGGACACTCAGGCAAGTCTTGATCACTGTGTATGTCAACAGAGGATACCGTCTCT 2826  
QY 2590 TGAAGAGAAACTGTCTCACTCTCTCATGCTGTCTAGCTCATACCCATTTCTTTTTC 2649  
Db 2827 TGAAGAGAAACTGTCTCACTCTCTCATGCTGTCTAGCTCATACCCATTTCTTTTTC 2886  
QY 2650 TTCACAGGTTTAACTGGTTTTTGGCATCTGCTATATATATCTCTCTCTCTCTCT 2709  
Db 2887 TTCACAGGTTTAACTGGTTTTTGGCATCTGCTATATATATCTCTCTCTCTCTCT 2946  
QY 2710 TATCTCTCCCTCCCTCCCTCCCTCTCTCTCCATCTCCATCTTTTGAATTTCTCTCAT 2769  
Db 2947 TATCTCTCCCTCCCTCCCTCCCTCTCTCTCCATCTCCATCTTTTGAATTTCTCTCAT 3006  
QY 2770 CCTTCATCTCAATCCCGTATCTAGCAACCCCGCCCCCCCCCAGGCAAGCAGTCTCTGA 2829  
Db 3007 CCTTCATCTCAATCCCGTATCTAGCAACCCCGCCCCCCCCCAGGCAAGCAGTCTCTGA 3065  
QY 2830 GTATCATATCACAAAGGACAAAGGCAACACACAAACCAGCCTCACTTACACTT 2889  
Db 3066 GTATCATATCACAAAGGACAAAGGCAACACACAAACCAGCCTCACTTACACTT 3125  
QY 2890 GGTACTCAAAAGAACAGAGTCAATGGTACTTGTCTAGCGTTTGGAGAGGAAACA 2949  
Db 3126 GGTACTCAAAAGAACAGAGTCAATGGTACTTGTCTAGCGTTTGGAGAGGAAACA 3185  
QY 2950 GGAACCCACCAACCAACCAATCAACCAACAAAGAAATTTCCAAATGAAGATGT 3009  
Db 3186 GGAACCCACCAACCAACCAATCAACCAACAAAGAAATTTCCAAATGAAGATGT 3245  
QY 3010 ATTTTGTCTTTTTCATTTTGGTGTATAGGCAATATATTCAGCAAAATGATTCCTTTC 3069  
Db 3246 ATTTTGTCTTTTTCATTTTGGTGTATAGGCAATATATTCAGCAAAATGATTCCTTTC 3305  
QY 3070 TTT-AAAAAATAATGTGAGGAAAGTAGAAATTTTACCAAGTTGTGGCCCGAGGGGT 3128  
Db 3306 TTTAAAAAATAATGTGAGGAAAGTAGAAATTTTACCAAGTTGTGGCCCGAGGGGT 3365

RESULT 8

AAS26153

ID AAS26153 standard; cDNA; 2290 BP.

XX AAS26153;

XX 07-NOV-2001 (first entry)

XX Human cDNA encoding a novel secreted protein, Seq ID 332.

XX Human; immunosuppressive; antiarthritic; ss; antirheumatic; cytostatic;  
KW cardiant; vasotropic; cerebroprotective; nootropic; neuroprotective;  
KW antibacterial; virucide; fungicide; ophthalmological; vulnerary;  
KW secreted protein; rheumatoid arthritis; hyperproliferative disorder;  
KW cardiovascular disorder; cardiac arrest; cerebrovascular disorder;  
KW cerebral ischaemia; angiogenesis; nervous system disorder;  
KW Alzheimer's disease; infection; ocular disorder; corneal infection;  
KW wound healing; epithelial cell proliferation; skin ageing; food additive;  
KW preservative; antiproliferative.

XX Homo sapiens.

XX WO200155322-A2.

XX 02-AUG-2001.

XX 17-JAN-2001; 2001WO-US001341.

XX 31-JAN-2000; 2000US-0179065P.

PR 04-FEB-2000; 2000US-0180628P.

PR 24-FEB-2000; 2000US-0184664P.

PR 02-MAR-2000; 2000US-0186350P.

PR 16-MAR-2000; 2000US-0189874P.

PR 17-MAR-2000; 2000US-0190076P.

PR 18-APR-2000; 2000US-0198123P.

PR 19-MAY-2000; 2000US-0205515P.

PR 07-JUN-2000; 2000US-0209467P.

PR 28-JUN-2000; 2000US-0214886P.

PR 30-JUN-2000; 2000US-0215135P.

PR 07-JUL-2000; 2000US-0216647P.

PR 07-JUL-2000; 2000US-0216880P.

PR 11-JUL-2000; 2000US-0217487P.

PR 11-JUL-2000; 2000US-0217496P.

PR 14-JUL-2000; 2000US-0218290P.

PR 26-JUL-2000; 2000US-0220963P.

PR 26-JUL-2000; 2000US-0220964P.

PR 14-AUG-2000; 2000US-0224518P.

PR 14-AUG-2000; 2000US-0224519P.

PR 14-AUG-2000; 2000US-0225213P.

PR 14-AUG-2000; 2000US-0225214P.

PR 14-AUG-2000; 2000US-0225266P.

PR 14-AUG-2000; 2000US-0225267P.

PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225447P.  
PR 14-AUG-2000; 2000US-0225477P.  
PR 14-AUG-2000; 2000US-0225757P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0226686P.  
PR 22-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 23-AUG-2000; 2000US-0228924P.  
PR 30-AUG-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 14-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 25-SEP-2000; 2000US-0234998P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235836P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239935P.  
PR 13-OCT-2000; 2000US-0239937P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241212P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 20-OCT-2000; 2000US-0241826P.  
PR 01-NOV-2000; 2000US-0241826P.  
PR 08-NOV-2000; 2000US-0246474P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.

PR 08-NOV-2000; 2000US-024527P.  
PR 08-NOV-2000; 2000US-024528P.  
PR 08-NOV-2000; 2000US-024532P.  
PR 08-NOV-2000; 2000US-024532P.  
PR 08-NOV-2000; 2000US-024609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249264P.  
PR 17-NOV-2000; 2000US-0249265P.  
PR 17-NOV-2000; 2000US-0249297P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 17-NOV-2000; 2000US-0249300P.  
PR 01-DEC-2000; 2000US-0250160P.  
PR 01-DEC-2000; 2000US-0250391P.  
PR 05-DEC-2000; 2000US-0251030P.  
PR 05-DEC-2000; 2000US-0251988P.  
PR 05-DEC-2000; 2000US-0256719P.  
PR 06-DEC-2000; 2000US-0251479P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.  
PR 08-DEC-2000; 2000US-0251989P.  
PR 08-DEC-2000; 2000US-0251990P.  
PR 11-DEC-2000; 2000US-0254097P.  
PR 05-JAN-2001; 2001US-0259678P.

(HUMA-) HUMAN GENOME SCI INC.  
Rosen CA, Barash SC, Ruben SM;  
WPI; 2001-488783/53.  
P-PSDB; AAU16166.

New nucleic acid molecules encoding 461 human secreted proteins for  
diagnosing, preventing, treating or ameliorating medical conditions and  
used as food additives or preservatives.

Claim 1; SEQ ID NO 332; 980pp; English.

The invention relates to isolated nucleic acid molecules and their  
encoded secreted proteins. The nucleic acids and proteins are used to  
prevent, treat or ameliorate a medical condition in e.g. humans, mice,  
rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used  
in diagnosing a pathological condition or susceptibility to a  
pathological condition. Antibodies to the proteins can also be used in  
alleviating symptoms associated with the disorders and in diagnostic  
immunoassays e.g. radioimmunoassays or enzyme linked immunosorbant assays  
(ELISA). Disorders which are diagnosed or treated include autoimmune  
diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g.  
neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac  
arrest, cerebrovascular disorders e.g. cerebral ischaemia, angiogenesis,  
nervous system disorders e.g. Alzheimer's disease, infections caused by  
bacteria, viruses and fungi and ocular disorders e.g. corneal infection,  
and many other disorders listed in the specification. The polypeptides  
can also be used to aid wound healing and epithelial cell proliferation,  
to prevent skin aging due to sunburn, to maintain organs before  
transplantation, for supporting cell culture of primary tissues, to  
regenerate tissues and in chemotaxis. The polypeptides can also be used  
as a food additive or preservative to increase or decrease storage







AC ABX73494;  
XX 18-MAR-2003 (first entry)  
XX Human novel polynucleotide #322.  
DE  
XX Human; gene; ds; neural disorder; immune system disorder; renal disorder;  
KW muscular disorder; respiratory disease; reproductive disorder;  
KW gastrointestinal disorder; pulmonary disorder; cardiovascular disorder;  
KW hyperproliferative disorder; inflammatory disease; allergic reaction;  
KW blood related disorder; cancer; immunosuppressive; antiinflammatory;  
KW cardiovascular; nephrotropic; cytostatic; antiallergic; thrombolytic;  
KW haemostatic; antiarteriosclerotic.  
XX  
OS Homo sapiens.  
XX  
XX US2002132753-A1.  
XX  
XX 19-SEP-2002.  
XX  
XX 17-JAN-2001; 2001US-00764864.  
XX  
XX 31-JAN-2000; 2000US-0179065P.  
PR 04-FEB-2000; 2000US-0180628P.  
PR 28-JUN-2000; 2000US-0214886P.  
PR 07-JUL-2000; 2000US-0216847P.  
PR 07-JUL-2000; 2000US-0216880P.  
PR 11-JUL-2000; 2000US-0217487P.  
PR 11-JUL-2000; 2000US-0217496P.  
PR 14-JUL-2000; 2000US-0218290P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 26-JUL-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225447P.  
PR 14-AUG-2000; 2000US-0225757P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 22-AUG-2000; 2000US-0226868P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 08-SEP-2000; 2000US-0229513P.  
PR 21-SEP-2000; 2000US-0231413P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234937P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239335P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.  
XX 08-DEC-2000; 2000US-0251869P.  
XX

PA (ROSE/) ROSEN C A.  
PA (RUBE/) RUBEN S M.  
XX (BARA/) BARASH S C.  
XX  
PI Rosen CA, Ruben SM, Barash SC;  
XX  
DR WPI; 2003-147444/14.  
DR P-PSDB; ABU55234.  
XX  
XX New polypeptides and nucleic acids, useful in gene therapy for treating,  
PT inhibiting or preventing e.g. neural, immune system, muscular,  
PT respiratory, reproductive, gastrointestinal, pulmonary, cardiovascular or  
PT renal disorders.  
XX  
PS Claim 1; SEQ ID NO 332; 402pp; English.  
XX  
XX The invention relates to human novel polypeptides and their associated  
CC polynucleotides. The polypeptides and polynucleotides are useful in gene  
CC therapy for treating, inhibiting or preventing neural disorders, immune  
CC system disorders (e.g. systemic lupus erythematosus, rheumatoid arthritis  
CC and multiple sclerosis), muscular disorders, respiratory diseases (e.g.  
CC nasal vestibulitis, nasal polyps and sinusitis), reproductive disorders,  
CC gastrointestinal disorders, pulmonary disorders, cardiovascular disorders  
CC (e.g. congenital heart defects, Ebstein's anomaly and hypoplastic left  
CC heart syndrome), renal disorders (e.g. acute kidney failure and end-stage  
CC renal disease), hyperproliferative disorders (e.g. Hodgkin's disease and  
CC leukaemia), inflammatory diseases (e.g. septic shock, bursitis and  
CC appendicitis), allergic reactions and conditions (e.g. asthma), blood  
CC related disorders (e.g. thrombosis, atherosclerosis and myocardial  
CC infarction) and cancerous diseases. Sequences ABX73173-ABX74167 represent  
CC human novel polynucleotides of the invention  
XX  
SQ Sequence 2290 BP; 707 A; 561 C; 491 G; 531 T; 0 U; 0 Other;  
  
Query Match 43.8%; Score 1496; DB 7; Length 2290;  
Best Local Similarity 99.7%; Pred. No. 0;  
Matches 1966; Conservative 0; Mismatches 1; Indels 4; Gaps 3;  
  
QY 1387 GCGCCATCATCGGGAGAGAGGGGGGCACATCAACAGCTGGCGAGATTGGCGGAGCC 1446  
DB 36 GCGCCATCATCGGGAGAGAGGGGGGCACATCAACAGCTGGCGAGATTGGCGGAGCC 95  
  
QY 1447 TCTATCAAGATTGCGCCCTCGGAAGGCCAGAGCTGAGCAAGAGTGTCTCATCATCACC 1506  
DB 96 TCTATCAAGATTGCGCCCTCGGAAGGCCAGAGCTGAGCAAGAGTGTCTCATCATCACC 155  
  
QY 1507 GGGCCACCGGAGCCCGAGTTCAAGGGCCAGGGACGGATCTTTGGGAACTGAAAGAGGAA 1566  
DB 156 GGGCCACCGGAGCCCGAGTTCAAGGGCCAGGGACGGATCTTTGGGAACTGAAAGAGGAA 215  
  
QY 1567 AACTTCTTTAACCCCAAGAGAGTGAAGCTGGAAGCGCATATCAGAGTGCCCTCTTCC 1626  
DB 216 AACTTCTTTAACCCCAAGAGAGTGAAGCTGGAAGCGCATATCAGAGTGCCCTCTTCC 275  
  
QY 1627 ACAGCTGGCGGGTGAATTGSCAAAGTGGCAACACCGTGAACCACTGCAAGACTTAACC 1686  
DB 276 ACAGCTGGCGGGTGAATTGSCAAAGTGGCAACACCGTGAACCACTGCAAGACTTAACC 335  
  
QY 1687 AGTGCAGAGTATCGTGCCCTCGTACCACCAACCCAGATGAAAATGAGGAAGTGATCGTC 1746  
DB 336 AGTGCAGAGTATCGTGCCCTCGTACCACCAACCCAGATGAAAATGAGGAAGTGATCGTC 395  
  
QY 1747 AGAATTATCGGGCACTTTCTTTGCTAGCCAGACTGCACAGCGCAAGATCAGGGAATTTGTA 1806  
DB 396 AGAATTATCGGGCACTTTCTTTGCTAGCCAGACTGCACAGCGCAAGATCAGGGAATTTGTA 455  
  
QY 1807 CAACAGGTGAAGCAGGAGCAGCAAGTAACCTCAGGGAGTCCCTCAGAGCCGCAAG 1866  
DB 456 CAACAGGTGAAGCAGGAGCAGCAAGTAACCTCAGGGAGTCCCTCAGAGCCGCAAG 515  
  
QY 1867 TGAGGGTCCCAACAGGACCAAGCAAAACACAGGATGAATGTAGCCCTTCCAAACCTGACA 1926  
DB 516 TGAGGGTCCCAACAGGACCAAGCAAAACACAGGATGAATGTAGCCCTTCCAAACCTGACA 575



CC detecting at least one differentially expressed gene product in a test  
CC sample derived from a cell suspected of being cancerous, where detection  
CC of the differentially expressed gene product is correlated with a  
CC cancerous state of the cell from which the test sample was derived. The  
CC polynucleotide sequences can be used in a method for detecting  
CC differentially expressed genes correlated with a cancerous state of a  
CC mammalian cell. The polynucleotides can also be used as probes for  
CC detecting and mapping related genes. They can be used in diagnosis and  
CC prognosis of diseases and disorders (e.g. identification of pre-  
CC metastatic or metastatic cancerous states, stages of cancer, or  
CC responsiveness of cancer to therapy). This is particularly for breast  
CC cancer, oestrogen receptor-positive breast cancer, oestrogen receptor-  
CC negative breast cancer, lung cancer, and colon cancer  
XX  
SQ Sequence 710 BP; 199 A; 174 C; 180 G; 152 T; 0 U; 5 Other;  
Query Match 16.6%; Score 567; DB 3; Length 710;  
Best Local Similarity 100.0%; Pred. No. 2.2e-243;  
Matches 567; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 534 CCGGATGAAGAGGTGAGCTCCCTTCGCCCTCAGCGAGCCAGCGTGGGACCACTC 593  
Db CCGGATGAAGAGGTGAGCTCCCTTCGCCCTCAGCGAGCCAGCGTGGGACCACTC 84  
QY 594 TTCCCGGAGCAAGCGCCCTCGGGGCACTTCTCAGGCGACAGATTGATTTCCTCC 653  
Db TTCCCGGAGCAAGCGCCCTCGGGGCACTTCTCAGGCGACAGATTGATTTCCTCC 144  
QY 654 GCTGCGGATCCTGGTCCCAACCCAGTTTGTGTGCTCATCGGAAAGGAGGCTTGAC 713  
Db GCTGCGGATCCTGGTCCCAACCCAGTTTGTGTGCTCATCGGAAAGGAGGCTTGAC 204  
QY 714 CATAAAGAACATCACTAAGCAGACCCAGTCCCGGTAGATATCCATAGAAAGAACTC 773  
Db CATAAAGAACATCACTAAGCAGACCCAGTCCCGGTAGATATCCATAGAAAGAACTC 264  
QY 774 TGAGCTGCAGAGAGCTGTCCATCCATGCCATGCCACCCAGAGGGGACTTCTGAAGCATG 833  
Db TGAGCTGCAGAGAGCTGTCCATCCATGCCATGCCACCCAGAGGGGACTTCTGAAGCATG 324  
QY 834 CCGCATGATCTTGAATCATGAGAAAGGAGGAGATGACCAACTAGCCGAGAGAT 893  
Db CCGCATGATCTTGAATCATGAGAAAGGAGGAGATGACCAACTAGCCGAGAGAT 384  
QY 894 TCCTCTGAAATCTTGCCACACAATGCTGTTGGTGAAGACTGATGGAAGAGGAGCAG 953  
Db TCCTCTGAAATCTTGCCACACAATGCTGTTGGTGAAGACTGATGGAAGAGGAGCAG 444  
QY 954 AAATTTGAAGAAATTGAACATGAAACAGGACCCAGATACCAATCTCATCTTTGCAGGA 1013  
Db AAATTTGAAGAAATTGAACATGAAACAGGACCCAGATACCAATCTCATCTTTGCAGGA 504  
QY 1014 TTTGAGCATATACAACCCGAAAGAACCACTGCTGAAAGGACACAGTTGAGGCTGTGC 1073  
Db TTTGAGCATATACAACCCGAAAGAACCACTGCTGAAAGGACACAGTTGAGGCTGTGC 564  
QY 1074 CAGTGTGAGATAGAGATTATGAAGAA 1100  
Db CAGTGTGAGATAGAGATTATGAAGAA 591

RESULT 11  
AAS26570  
ID AAS26570 standard; cDNA; 1186 BP.  
XX AC AAS26570;  
XX DT 07-NOV-2001 (first entry)  
XX Human cDNA encoding a novel secreted protein, seq ID 749.  
DE Human; immunosuppressive; antiarthritic; ss; antirheumatic; cytostatic;  
KW Human; vasotropic; cerebroprotective; nootropic; neuroprotective;  
KW

KW antibacterial; virucide; fungicide; ophthalmological; vulnery;  
KW secreted protein; rheumatoid arthritis; hyperproliferative disorder;  
KW cardiovascular disorder; cardiac arrest; cerebrovascular disorder;  
KW cerebral ischaemia; angiogenesis; nervous system disorder;  
KW Alzheimer's disease; infection; ocular disorder; corneal infection;  
KW wound healing; epithelial cell proliferation; skin ageing; food additive;  
KW preservative; antiproliferative.  
XX Homo sapiens.  
XX WO200155322-A2.  
PD 02-AUG-2001.  
XX 17-JAN-2001; 2001WO-US001341.  
XX 31-JAN-2000; 2000US-0179065P.  
PR 04-FEB-2000; 2000US-0180628P.  
PR 24-FEB-2000; 2000US-0184664P.  
PR 02-MAR-2000; 2000US-0186350P.  
PR 16-MAR-2000; 2000US-0189874P.  
PR 17-MAR-2000; 2000US-0190076P.  
PR 18-APR-2000; 2000US-0198123P.  
PR 19-MAY-2000; 2000US-0205515P.  
PR 07-JUN-2000; 2000US-0209467P.  
PR 28-JUN-2000; 2000US-0214886P.  
PR 30-JUN-2000; 2000US-0215135P.  
PR 07-JUL-2000; 2000US-0216880P.  
PR 11-JUL-2000; 2000US-0217487P.  
PR 14-JUL-2000; 2000US-0217496P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 26-JUL-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225213P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225266P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225447P.  
PR 14-AUG-2000; 2000US-0225757P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0226688P.  
PR 22-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.

PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234224P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 25-SEP-2000; 2000US-0234998P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235836P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239935P.  
PR 13-OCT-2000; 2000US-0239937P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 20-OCT-2000; 2000US-0241826P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0246474P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249219P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249246P.  
PR 17-NOV-2000; 2000US-0249264P.  
PR 17-NOV-2000; 2000US-0249265P.  
PR 17-NOV-2000; 2000US-0249266P.  
PR 17-NOV-2000; 2000US-0249297P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 01-DEC-2000; 2000US-0250300P.  
PR 01-DEC-2000; 2000US-0250316P.  
PR 05-DEC-2000; 2000US-0250391P.  
PR 05-DEC-2000; 2000US-0251030P.  
PR 05-DEC-2000; 2000US-0251988P.  
PR 06-DEC-2000; 2000US-0256719P.  
PR 08-DEC-2000; 2000US-0251479P.  
PR 08-DEC-2000; 2000US-0251856P.

PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.  
PR 08-DEC-2000; 2000US-0251989P.  
PR 08-DEC-2000; 2000US-0251990P.  
PR 11-DEC-2000; 2000US-0254097P.  
PR 05-JAN-2001; 2001US-0259678P.  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX  
PI Rosen CA, Barash SC, Ruben SM;  
DR WPI; 2001-488783/53.  
DR P-PSDB; AAU16583.  
XX  
XX  
PT New nucleic acid molecules encoding 461 human secreted proteins for  
PT diagnosing, preventing, treating or ameliorating medical conditions and  
PT used as food additives or preservatives.  
XX  
PS Claim 1; SEQ ID NO 749; 980pp; English.  
XX  
CC The invention relates to isolated nucleic acid molecules and their  
CC encoded secreted proteins. The nucleic acids and proteins are used to  
CC prevent, treat or ameliorate a medical condition in e.g. humans, mice,  
CC rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used  
CC in diagnosing a pathological condition or susceptibility to a  
CC pathological condition. Antibodies to the proteins can also be used in  
CC alleviating symptoms associated with the disorders and in diagnostic  
CC immunoassays e.g. radioimmunoassays or enzyme linked immunosorbent assays  
CC (ELISA). Disorders which are diagnosed or treated include autoimmune  
CC diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g.  
CC neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac  
CC arrest, cerebrovascular disorders e.g. cerebral ischaemia, angiodysplasia,  
CC nervous system disorders e.g. Alzheimer's disease, infections caused by  
CC bacteria, viruses and fungi and ocular disorders e.g. corneal infection,  
CC and many other disorders listed in the specification. The polypeptides  
CC can also be used to aid wound healing and epithelial cell proliferation,  
CC to prevent skin aging due to sunburn, to maintain organs before  
CC transplantation, for supporting cell culture of primary tissues, to  
CC regenerate tissues and in chemotaxis. The polypeptides can also be used  
CC as a food additive or preservative to increase or decrease storage  
CC capabilities, fat content, lipid, protein, carbohydrate, vitamins,  
CC minerals, cofactors and other nutritional components. The present  
CC sequence encodes a novel secreted protein of the invention. Note: The  
CC sequence data for this patent did not form part of the printed

Query Match 11.8%; Score 401; DB 4; Length 1186;

Best Local Similarity 99.4%; Pred. No. 5.1e-169;

Matches 671; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 1864 AAGTGGGCTCCACAGGACCCAGCAGCAACCAACGAGTGAATGTAGCCCTTCCAACTG 1923

Db 512 AAGTGGGCTCCACAGGACCCAGCAGCAACCAACGAGTGAATGTAGCCCTTCCAACTG 571

QY 1924 ACAGATGAGACCAACGACGAGCCAGATCGGAGCAACCAACGACCATCTGAGGAA 1983

Db 572 ACAGATGAGACCAACGACGAGCCAGATCGGAGCAACCAACGACCATCTGAGGAA 631

QY 1984 TGAGAGTCTGCGGAGCGGCCAGGAGCTCTGCGAGGCGCTCGAGAACCCAGGCGCGGA 2043

Db 632 TGAGAGTCTGCGGAGCGGCCAGGAGCTCTGCGAGGCGCTCGAGAACCCAGGCGCGGA 691

QY 2044 GGAGGGCGGGGAAGGTTCAGCCAGGTTTGCCAGAACCCAGGCGCGCTCCGCGCC 2103

Db 692 GGAGGGCGGGGAAGGTTCAGCCAGGTTTGCCAGAACCCAGGCGCGCTCCGCGCC 751

QY 2104 CAGGGCTTTCAGAGGCTTCAGCCAT-CCACTTCACCATCCACTCGGATCTCTCTGAAC 2162

Db 752 CAGGGCTTTCAGAGGCTTCAGCNATNCCACTTCACCATCCACTCGGATCTCTCTGAAC 811

QY 2163 CCCACGACGTATCCCTTTTAGTTGAACATAAGTAGTGAACGTGTTCGAAGCCAGCAA 2222

Db 812 CCCACGACGTATCCCTTTTAGTTGAACATAAGTAGTGAACGTGTTCGAAGCCAGCAA 871

QY 2223 AATGCACACCCCTTTTCTGTGGCAATCGTCTGTACATGTGTGTACATATAGAAAG 2282  
DB 872 AATGCACACCCCTTTTCTGTGGCAATCGTCTGTACATGTGTGTACATATAGAAAG 931  
QY 2283 GAAGATGTTAAGATATGTGGCTGTGGTTACACAGGTCCTGCAGCGGTAATATATTT 2342  
DB 932 GAAGATGTTAAGATATGTGGCTGTGGTTACACAGGTCCTGCAGCGGTAATATATTT 991  
QY 2343 TAGAATAATATATCAATACTCAACTCAACTCAACTCAACTCAACTCAACTCAACTCAACT 2402  
DB 992 TAGAATAATATATCAATACTCAACTCAACTCAACTCAACTCAACTCAACTCAACTCAACT 1051  
QY 2403 TTCTCTTTTAAAG 2462  
DB 1052 TTCTCTTTTAAAG 1111  
QY 2463 CTCACGGTGTAG 2522  
DB 1112 CTCACGGTGTAG 1171  
QY 2523 CGTCGGAAGGACACT 2537  
DB 1172 CGTCGGAAGGACACT 1186  
RESULT 12  
ABX73911  
ID ABX73911 standard; DNA; 1186 BP.  
XX  
AC ABX73911;  
XX  
DT 18-MAR-2003 (first entry)  
XX  
DE Human novel polynucleotide #739.  
XX  
DE Human; gene; ds; neural disorder; immune system disorder; renal disorder;  
KW muscular disorder; respiratory disease; reproductive disorder;  
KW gastrointestinal disorder; pulmonary disorder; cardiovascular disorder;  
KW hyperproliferative disorder; inflammatory disease; allergic reaction;  
KW blood related disorder; cancer; immunosuppressive; antiinflammatory;  
KW cardiovascular; nephrotropic; cytostatic; antiallergic; thrombolytic;  
KW haemostatic; antiarteriosclerotic.  
XX  
OS Homo sapiens.  
XX  
PN US2002132753-A1.  
XX  
PD 19-SEP-2002.  
XX  
PF 17-JAN-2001; 2001US-00764864.  
XX  
PR 31-JAN-2000; 2000US-0179065P.  
PR 04-FEB-2000; 2000US-0180628P.  
PR 28-JUN-2000; 2000US-0214896P.  
PR 07-JUL-2000; 2000US-0216647P.  
PR 07-JUL-2000; 2000US-0216880P.  
PR 11-JUL-2000; 2000US-0217497P.  
PR 11-JUL-2000; 2000US-0218230P.  
PR 14-JUL-2000; 2000US-0219696P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 26-JUL-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225447P.  
PR 14-AUG-2000; 2000US-0225757P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 22-AUG-2000; 2000US-0226868P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.

PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239935P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.  
XX  
PA (ROSE/) ROSEN C A.  
PA (RUBE/) RUBEN S M.  
PA (BARA/) BARASH S C.  
XX  
PI Rosen CA, Ruben SM, Barash SC;  
XX  
PI WPI; 2003-147444/14.  
XX  
DR P-PSDB; ABUS5651.  
XX  
PT New polypeptides and nucleic acids, useful in gene therapy for treating,  
PT inhibiting or preventing e.g. neural, immune system, muscular,  
PT respiratory, reproductive, gastrointestinal, pulmonary, cardiovascular or  
PT renal disorders.  
XX  
PS Claim 1; SEQ ID NO 749; 402pp; English.  
XX  
CC The invention relates to human novel polypeptides and their associated  
CC polynucleotides. The polypeptides and polynucleotides are useful in gene  
CC therapy for treating, inhibiting or preventing neural disorders, immune  
CC system disorders (e.g. systemic lupus erythematosus, rheumatoid arthritis  
CC and multiple sclerosis), muscular disorders, respiratory diseases (e.g.  
CC nasal vestibulitis, nasal polyps and sinusitis), reproductive disorders,  
CC gastrointestinal disorders, pulmonary disorders, cardiovascular disorders  
CC (e.g. congenital heart defects, Ebstein's anomaly and hypoplastic left  
CC heart syndrome), renal disorders (e.g. acute kidney failure and end-stage  
CC renal disease), hyperproliferative disorders (e.g. Hodgkin's disease and  
CC leukaemia), inflammatory diseases (e.g. septic shock, bursitis and  
CC appendicitis), allergic reactions and conditions (e.g. asthma), blood  
CC related disorders (e.g. thrombosis, atherosclerosis and myocardial  
CC infarction) and cancerous diseases. Sequences ABX73173-ABX74167 represent  
CC human novel polynucleotides of the invention  
XX  
SQ Sequence 1186 BP; 338 A; 310 C; 302 G; 230 T; 0 U; 6 Other;  
Query Match 11.8%; Score 401; DB 7; Length 1186;  
Best Local Similarity 99.4%; Pred. No. 5.1e-169;  
Matches 671; Conservative 0; Mismatches 3; Indels 1; Gaps 1;  
QY 1864 AGTGAGGCTCCACAGGACAGGACAAACACGATGATGTAGCCCTTCCACACCTG 1923  
DB 512 AGTGAGGCTCCACAGGACAGGACAAACACGATGATGTAGCCCTTCCACACCTG 571  
QY 1924 ACAGATGAGACCAACAGCGAGCCAGATCGGGAGCAACCAAGACCATCTGAGGAA 1983









Search completed: July 21, 2004, 17:06:27  
Job time : 1245 secs

```
XX Human secreted protein 5' EST, SEQ ID NO: 3265.
DE
XX Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW gene therapy; chromosome mapping; ss.
XX Homo sapiens.
OS
XX EP1033401-A2.
PN
XX 06-SEP-2000.
PD
XX 21-FEB-2000; 2000EP-00200610.
PF
XX 26-FEB-1999; 99US-0122487P.
PR
XX (GEST ) GENSET.
PA
XX Dumas Milne Edwards J, Duclert A, Giordano J;
PI
XX WPI; 2000-500381/45.
DR
XX P-PSDB; AAG03261.
DR
XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for
PT diagnostic, forensic, gene therapy and chromosome mapping procedures.
XX
XX Claim 1; SEQ ID NO 3265; 71pp + Sequence Listing; English.
PS
XX
XX The present sequence is one of a large number of 5' ESTs derived from
CC mRNAs encoding secreted proteins. An ORP has been identified within the
CC sequence. The 5' ESTs were prepared from total human RNAs or polyA+ RNAs
CC derived from 30 different tissues. EST sequences usually correspond
CC mainly to the 3' untranslated region (UTR) of the mRNA because they are
CC often obtained from oligo-dT primed cDNA libraries. Such ESTs are not
CC well suited for isolating cDNA sequences derived from the 5' ends of
CC mRNAs and even in those cases where longer cDNA sequences have been
CC obtained, the full 5' UTR is rarely included. 5' ESTs are derived from
CC mRNAs with intact 5' ends and can therefore be used to obtain full length
CC cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic,
CC gene therapy and chromosome mapping procedures. They are used to obtain
CC upstream regulatory sequences and to design expression and secretion
CC vectors
XX
XX Sequence 364 BP; 84 A; 101 C; 118 G; 59 T; 0 U; 2 Other;
SQ
Query Match 10.2%; Score 347; DB 3; Length 364;
Best Local Similarity 100.0%; Pred. No. 8.1e-145;
Matches 347; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 15 CGAGGAGCGCGGGTACCGGGCCGGGGAGCGCGGGCTCTCGGGGAAGAGACGGATGAT 74
Db 18 CGAGGAGCGCGGGTACCGGGCCGGGGAGCGCGGGCTCTCGGGGAAGAGACGGATGAT 77
Qy 75 GACCAAGCTTTACATCGGACCTGAGCCCGCGCGTCCACCGCGACGACCTCCGGCAGCT 134
Db 78 GACCAAGCTTTACATCGGACCTGAGCCCGCGCGTCCACCGCGACGACCTCCGGCAGCT 137
Qy 135 CTTTGGGGACAGGAAGCTGCCCTCGGGGACAGGTCTCTGTAAGTCCGGCTACGCCTT 194
Db 138 CTTTGGGGACAGGAAGCTGCCCTCGGGGACAGGTCTCTGTAAGTCCGGCTACGCCTT 197
Qy 195 CGTGGACTACCCCGACAGAACTGGGCCATCCGGCCATCGAGACCCCTCTCGGGTAAAGT 254
Db 198 CGTGGACTACCCCGACAGAACTGGGCCATCCGGCCATCGAGACCCCTCTCGGGTAAAGT 257
Qy 255 GGAATTGCATCGGAAATCATGGAAGTTGATTACTCAGTCTCTAAAGCTAAGAGCAG 314
Db 258 GGAATTGCATCGGAAATCATGGAAGTTGATTACTCAGTCTCTAAAGCTAAGAGCAG 317
Qy 315 GAAATTTCAGATTGAAACATCCCTCCCTCAGTGGAGGTGT 361
Db 318 GAAATTTCAGATTGAAACATCCCTCCCTCAGTGGAGGTGT 364
```

**This Page Blank (uspto)**